This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

= EP1261932

(19) World Intellectual Property Organization International Bureau



<u>idara dibandi ku birdik ratidi biri kuko kirdi diban katal kalik kan dibaka kirdi dibi</u>

(43) International Publication Date 19 April 2001 (19.04.2001)

PCT

(10) International Publication Number WO 01/27857 A2

(51) International Pa	tent Classification7:	G06F 19/00	DEN BOOM, Dirk [DE/DE]; Eppendorfer Weg 205 D, D-20253 Hamburg (DE). PING, Yip [US/US]; 3641 Copley Avenue, San Diego, CA 92116 (US). RODI, Charlie [US/US]; 13823 Recuerdo Drive, Del Mar, CA	
(21) International Ap	plication Number:	PCT/US00/28413		
(22) International Filing Date: 13 October 2000 (13.10.2000)		2000 (13.10.2000)	92014 (US). HE, Liyan [CN/US]; 10948 Creek Bridge Place, San Diego, CA 92128 (US). CHIU, Norman	
(25) Filing Language	:	English	[CA/US]; 1128 Caminito Alvarez, San Diego, CA 92126 (US). JURINKE, Christian [DE/DE]; Rombergstrasse, 22, 20255 Hamburg (DE).	
(26) Publication Language: English		English	22, 2023 Hamburg (32).	
(30) Priority Data: 60/159,176	30) Priority Data:		(74) Agents: SEIDMAN, Stephanie, L. et al.; Heller Ehrman White & McAuliffe, Suite 700, 4250 Executive Square, La Jolla, CA 92037 (US).	

US

US

US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

10 July 2000 (10.07.2000)

10 July 2000 (10.07.2000)

19 September 2000 (19.09.2000)

60/159,176 (CIP) US 13 October 1999 (13.10.1999) Filed on 60/217,658 (CIP) US 10 July 2000 (10.07.2000) Filed on 09/663,968 (CIP) US 19 September 2000 (19.09.2000) Filed on 60/217.251 (CIP) US 10 July 2000 (10.07.2000) Filed on

- (71) Applicant (for all designated States except US): SE-QUENOM, INC. [US/US]; 11555 Sorrento Valley Road, San Diego, CA 92121-1331 (US).
- (72) Inventors; and

60/217,658

60/217,251

09/663,968

(75) Inventors/Applicants (for US only): BRAUN, Andreas [DE/US]; 11237-6 Carmel Creek Road, San Diego, CA 92130 (US). KÖSTER, Hubert [DE/US]; 8636-C Via Mallorca Drive, La Jolla, CA 92037 (US). VAN

- Jolla, CA 92037 (US).
- (81) Designated States (national): AE, AG, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS

(57) Abstract: Process and methods for creating a database of genomic samples from healthy human donors, methods that use the database to identify and correlate polymorphic genetic markers and other markers with diseases and conditions are provided.

METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS

RELATED APPLICATIONS

Benefit of priority to the following applications is claimed herein:

U.S. provisional application Serial No. 60/217,658 to Andreas Braun, Hubert Koster; Dirk Van den Boom, filed July 10, entitled "METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS"; U.S. provisional application Serial No. 60/159,176 to Andreas Braun, Hubert Koster, Dirk Van den Boom, filed October 13, 1999, entitled "METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS"; U.S. provisional application Serial No. 60/217,251, filed July 10, 2000, to Andreas Braun, entitled "POLYMORPHIC KINASE ANCHOR PROTEIN GENE SEQUENCES, POLYMORPHIC KINASE ANCHOR PROTEINS AND METHODS OF DETECTING POLYMORPHIC KINASE ANCHOR PROTEINS AND NUCLEIC ACIDS ENCODING THE SAME"; and U.S. application Serial No. 09/663,968, to Ping Yip, filed September 19, 2000, entitled "METHOD AND DEVICE FOR IDENTIFYING A BIOLOGICAL SAMPLE."

Where permitted that above-noted applications and provisional applications are incorporated by reference in their entirety.

FIELD OF THE INVENTION

Process and methods for creating a database of genomic samples from healthy human donors. Methods that use the database to identify and correlate with polymorphic genetic markers and other markers with diseases and conditions are provided.

BACKGROUND

25

Diseases in all organisms have a genetic component, whether inherited or resulting from the body's response to environmental stresses, such as viruses and toxins. The ultimate goal of ongoing genomic research is to use this information to develop new ways to identify, treat and potentially cure these diseases. The first step has been to screen disease tissue and identify genomic changes at the level of individual samples. The identification of these "disease"

markers has then fueled the development and commercialization of diagnostic tests that detect these errant genes or polymorphisms. With the increasing numbers of genetic markers, including single nucleotide polymorphisms (SNPs), microsatellites, tandem repeats, newly mapped introns and exons, the challenge to the medical and pharmaceutical communities is to identify genotypes which not only identify the disease but also follow the progression of the disease and are predictive of an organism's response to treatment.

Currently the pharmaceutical and biotechnology industries find a disease and then attempt to determine the genomic basis for the disease. This approach is time consuming and expensive and in many cases involves the investigator guessing as to what pathways might be involved in the disease.

Genomics

Presently the two main strategies employed in analyzing the available genomic information are the technology driven reverse genetics brute force strategy and the knowledge-based pathway oriented forward genetics strategy. The brute force approach yields large databases of sequence information but little information about the medical or other uses of the sequence information. Hence this strategy yields intangible products of questionable value. The knowledge-based strategy yields small databases that contain a lot of information about medical uses of particular DNA sequences and other products in the pathway and yield tangible products with a high value.

Polymorphisms

25

30

Polymorphisms have been known since 1901 with the identification of blood types. In the 1950's they were identified on the level of proteins using large population genetic studies. In the 1980's and 1990's many of the known protein polymorphisms were correlated with genetic loci on genomic DNA. For example, the gene dose of the apolipoprotein E type 4 allele was correlated with the risk of Alzheimer's disease in late onset families (see, e.g., Corder et al. (1993) Science 261: 921-923; mutation in blood coagulation factor V was associated with resistance to activated protein C (see, e.g., Bertina et al. (1994) Nature 369:64-67); resistance to HIV-1 infection has been shown in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene (see,

PCT/US00/28413

10

15

30

e.g., Samson et al. (1996) Nature 382:722-725); and a hypermutable tract in antigen presenting cells (APC, such as macrophages), has been identified in familial colorectal cancer in individuals of Ashkenzi jewish background (see, e.g., Laken et al. (1997) Nature Genet. 17:79-83). There may be more than three million polymorphic sites in the human genome. Many have been identified, but not yet characterized or mapped or associated with a marker.

Single nucleotide polymorphisms (SNPs)

Much of the focus of genomics has been in the identification of SNPs, which are important for a variety of reasons. They allow indirect testing (association of haplotypes) and direct testing (functional variants). They are the most abundant and stable genetic markers. Common diseases are best explained by common genetic alterations, and the natural variation in the human population aids in understanding disease, therapy and environmental interactions.

Currently, the only available method to identify SNPs in DNA is by sequencing, which is expensive, difficult and laborious. Furthermore, once a SNP is discovered it must be validated to determine if it is a real polymorphism and not a sequencing error. Also, discovered SNPs must then be evaluated to determine if they are associated with a particular phenotype. Thus, there is a 20 need to develop new paradigms for identifying the genomic basis for disease and markers thereof. Therefore, it is an object herein to provide methods for identifying the genomic basis of disease and markers thereof.

SUMMARY

Databases and methods using the databases are provided herein. The databases comprise sets of parameters associated with subjects in populations selected only on the basis of being healthy (i.e., where the subjects are mammals, such as humans, they are selected based upon apparent health and no detectable infections). The databases can be sorted based upon one or more of the selected parameters.

The databases are preferably relational databases, in which an index that represents each subject serves to relate parameters, which are the data, such as age, ethnicity, sex, medical history, etc. and ultimately genotypic information,

PCT/US00/28413

15

20

30

that was inputted into and stored in the database. The database can then be sorted according to these parameters. Initially, the parameter information is obtained from a questionnaire answered by each subject from whom a body tissue or body fluid sample is obtained. As additional information about each sample is obtained, this information can be entered into the database and can serve as a sorting parameter.

The databases obtained from healthy individuals have numerous uses, such as correlating known polymorphisms with a phenotype or disease. The databases can be used to identify alleles that are deleterious, that are beneficial, and that are correlated with diseases.

For purposes herein, genotypic information can be obtained by any method known to those of skill in the art, but is preferably obtained using mass spectrometry.

Also provided herein, is a new use for existing databases of subjects and genotypic and other parameters, such as age, ethnicity, race, and gender. Any database can be sorted according to the methods herein and alleles that exhibit statistically significant correlations with any of the sorting parameters can be identified. It is noted, however, is noted, that the databases provided herein and randomly selected databases will perform better in these methods, since diseasebased databases suffer numerous limitations, including their relatively small size, the homogeneity of the selected disease population, and the masking effect of the polymorphism associated with the markers for which the database was selected. Hence, the healthy database provided herein, provides advantages not heretofore recognized or exploited. However, the methods provided herein can 25 be used with a selected database, including disease-based databases, with or without sorting for the discovery and correlation of polymorphisms. In addition, the databases provided herein represent a greater genetic diversity than the unselected databases typically utilized for the discovery of polymorphisms and thus allow for the enhanced discovery and correlation of polymorphisms.

The databases provided herein can be used for taking an identified polymorphism, and ascertaining whether it changes in frequency when the data is sorted according to a selected parameter.

20

25

30

treatment regimens.

One use of these methods is correlating a selected marker with a particular parameter by following the occurrence of known genetic markers and then, having made this correlation, determining or identifying correlations with diseases. Examples of this use are p53 and Lipoprotein Lipase polymorphism.

As exemplified herein, known markers are shown to have particular correlation with certain groups, such as a particular ethnicity or race or one sex. Such correlations will then permit development of better diagnostic tests and

These methods are valuable for identifying one or more genetic markers whose frequency changes within the population as a function of age, ethnic group, sex or some other criteria. This can allow the identification of previously unknown polymorphisms and ultimately a gene or pathway involved in the onset and progression of disease.

The databases and methods provided herein permit, among other things, identification of components, particularly key components, of a disease process by understanding its genetic underpinnings and also permit an understanding of processes, such as individual drug responses. The databases and methods provided herein also can be used in methods involving elucidation of pathological pathways, in developing new diagnostic assays, identifying new potential drug targets, and in identifying new drug candidates.

The methods and databases can be used with experimental procedures, including, but are not limited to, in silico SNP identification, in vitro SNP identification/verification, genetic profiling of large populations, and in biostatistical analyses and interpretations.

Also provided herein, are combinations that contain a database provided herein and a biological sample from a subject in the database, and preferably biological samples from all subjects or a plurality of subjects in the database. Collections of the tissue and body fluid samples are also provided.

Also, provided herein, are methods for determining a genetic marker that correlates with age, comprising identifying a polymorphism and determining the frequency of the polymorphism with increasing age in a healthy population.

PCT/US00/28413

5

10

20

Further provided herein are methods for determining whether a genetic marker correlates with susceptibility to morbidity, early mortality, or morbidity and early mortality, comprising identifying a polymorphism and determining the frequency of the polymorphism with increasing age in a healthy population.

Any of the methods herein described can be used out in a multiplex format.

Also provided are an apparatus and process for accurately identifying genetic information. It is another object of the herein that genetic information be extracted from genetic data in a highly automated manner. Therefore, to overcome the deficiencies in the known conventional systems, a method and apparatus for identifying a biological sample is proposed.

Briefly, the method and system for identifying a biological sample generates a data set indicative of the composition of the biological sample. In a particular example, the data set is DNA spectrometry data received from a mass spectrometer. The data set is denoised, and a baseline is deleted. Since possible compositions of the biological sample may be known, expected peak areas may be determined. Using the expected peak areas, a residual baseline is generated to further correct the data set. Probable peaks are then identifiable in the corrected data set, which are used to identify the composition of the biological sample. In a disclosed example, statistical methods are employed to determine the probability that a probable peak is an actual peak, not an actual peak, or that the data too inconclusive to call.

Advantageously, the method and system for identifying a biological sample accurately makes composition calls in a highly automated manner. In such a manner, complete SNP profile information, for example, may be collected efficiently. More importantly, the collected data is analyzed with highly accurate results. For example, when a particular composition is called, the result may be relied upon with great confidence. Such confidence is provided by the robust computational process employed.

30 DESCRIPTION OF THE DRAWINGS

15

20

25

Figure 1 depicts an exemplary sample bank. Panel 1 shows the samples as a function of sex and ethnicity. Panel 2 shows the caucasians as a function of age. Panel 3 shows the Hispanics as a function of age.

Figures 2A and 2C show an age- and sex-distribution of the 291S allele of the lipoprotein lipase gene in which a total of 436 males and 589 females were investigated. Figure 2B shows an age distribution for the 436 males.

Figure 3 is an exemplary questionnaire for population-based sample banking.

Figure 4 depicts processing and tracking of blood sample components.

Figure 5 depicts the allelic frequency of "sick" alleles and "healthy" alleles as a function of age. It is noted that the relative frequency of healthy alleles increases in a population with increasing age.

Figure 6 depicts the age-dependent distribution of ApoE genotypes (see, Schächter et al. (1994) Nature Genetics 6:29-32).

Figure 7A-D depicts age-related and genotype frequency of the p53 (tumor suppressor) codon 72 among the caucasian population in the database.

*B72 and *P72 represent the frequency of the allele in the database population.

R72, R72P, and P72 represent the genotypes of the individuals in the population.

The frequency of the homozygous P72 allele drops from 6.7% to 3.7% with age.

Figure 8 depicts the allele and genotype frequencies of the p21 S31R allele as a function of age.

Figure 9 depicts the frequency of the FVII Allele 353Q in pooled versus individual samples.

Figure 10 depicts the frequency of the CETP (cholesterol ester transfer protein) allele in pooled versus individual samples

Figure 11 depicts the frequency of the plasminogen activator inhibitor-1 (PAI-1) 5G in pooled versus individual samples

Figure 12 shows mass spectra of the samples and the ethnic diversity of 30 the PAI-1 alleles.

Figure 13 shows mass spectra of the samples and the ethnic diversity of the CETP 405 alleles.

10

20

30

Figure 14 shows mass spectra of the samples and the ethnic diversity of the Factor VII 353 alleles.

Figure 15 shows ethnic diversity of PAI-1, CETP and Factor VII using the pooled DNA samples.

Figure 16 shows the p53-Rb pathway and the relationships among the various factors in the pathway.

Figure 17, which is a block diagram of a computer constructed to provide and process the databases described herein, depicts a typical computer system for storing and sorting the databases provided herein and practicing the methods provided herein.

Figure 18 is a flow diagram that illustrates the processing steps performed using the computer illustrated in Figure 17, to maintain and provide access to the databases for identifying polymorphic genetic markers.

Figure 19 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the AKAP10-1 locus. Bright green bars show frequencies in individuals younger than 40 years. Dark green bars show frequencies in individuals older than 60 years.

Figure 20 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the AKAP10-5 locus. Bright green bars show frequencies in individuals younger than 40 years; dark green bars show frequencies in individuals older than 60 years.

Figure 21 is a histogram showing the allele and genotype distribution in the age and sex stratified Caucasian population for the h-msrA locus. Genotype difference between male age groups is significant. Bright green bars show frequencies in individuals younger than 40 years. Dark green bars show frequencies in individuals older than 60 years.

Figure 22A-D is a sample data collection questionnaire used for the healthy database.

Figure 23 is a flowchart showing processing performed by the computing device of Figure 24 when performing genotyping of sense strands and antisense strands from assay fragments.

. 5

15

20

Figure 24 is a block diagram showing a system in accordance with the present invention;

Figure 25 is a flowchart of a method of identifying a biological sample in accordance with the present invention;

Figure 26 is a graphical representation of data from a mass spectrometer;

Figure 27 is a diagram of wavelet transformation of mass spectrometry data;

Figure 28 is a graphical representation of wavelet stage 0 hi data;

Figure 29 is a graphical representation of stage 0 noise profile;

Figure 30 is a graphical representation of generating stage noise standard deviations;

Figure 31 is a graphical representation of applying a threshold to data stages;

Figure 32 is a graphical representation of a sparse data set;

Figure 33 is a formula for signal shifting;

Figure 34 is a graphical representation of a wavelet transformation of a denoised and shifted signal;

Figure 35 is a graphical representation of a denoised and shifted signal;

Figure 36 is a graphical representation of removing peak sections;

Figure 37 is a graphical representation of generating a peak free signal;

Figure 38 is a block diagram of a method of generating a baseline correction;

Figure 39 is a graphical representation of a baseline and signal;

Figure 40 is a graphical representation of a signal with baseline removed;

25 Figure 41 is a table showing compressed data;

Figure 42 is a flowchart of method for compressing data;

Figure 43 is a graphical representation of mass shifting;

Figure 44 is a graphical representation of determining peak width;

Figure 45 is a graphical representation of removing peaks;

30 Figure 46 is a graphical representation of a signal with peaks removed;

Figure 47 is a graphical representation of a residual baseline;

Figure 48 is a graphical representation of a signal with residual baseline removed;

Figure 49 is a graphical representation of determining peak height;

Figure 50 is a graphical representation of determining signal-to-noise for each peak;

Figure 51 is a graphical representation of determining a residual error for each peak;

Figure 52 is a graphical representation of peak probabilities;

Figure 53 is a graphical representation of applying an allelic ratio to peak probability;

Figure 54 is a graphical representation of determining peak probability Figure 55 is a graphical representation of calling a genotype;

Figure 56 is a flowchart showing a statistical procedure for calling a genotype;

Figure 57 is a flowchart showing processing performed by the computing device of Figure 1 when performing standardless genotyping; and

Figure 58 is graphical representation of applying an allelic ratio to peak probability for standardless genotype processing.

DETAILED DESCRIPTION

20 Definitions

25

30

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of ordinary skill in the art to which this invention belongs. All patents, applications, published applications and other publications and sequences from GenBank and other databases referred to herein throughout the disclosure are incorporated by reference in their entirety.

As used herein, a biopolymer includes, but is not limited to, nucleic acid, proteins, polysaccharides, lipids and other macromolecules. Nucleic acids include DNA, RNA, and fragments thereof. Nucleic acids may be derived from genomic DNA, RNA, mitochondrial nucleic acid, chloroplast nucleic acid and other organelles with separate genetic material.

PCT/US00/28413

15

20

25

30

-11-

As used herein, morbidity refers to conditions, such as diseases or disorders, that compromise the health and well-being of an organism, such as an animal. Morbidity susceptibility or morbidity-associated genes are genes that, when altered, for example, by a variation in nucleotide sequence, facilitate the expression of a specific disease clinical phenotype. Thus, morbidity susceptibility genes have the potential, upon alteration, of increasing the likelihood or general risk that an organism will develop a specific disease.

As used herein, mortality refers to the statistical likelihood that an organism, particularly an animal, will not survive a full predicted lifespan. Hence, a trait or a marker, such as a polymorphism, associated with increased mortality is observed at a lower frequency in older than younger segments of a population.

As used herein, a polymorphism, e.g. genetic variation, refers to a variation in the sequence of a gene in the genome amongst a population, such as allelic variations and other variations that arise or are observed. Thus, a polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. These differences can occur in coding and non-coding portions of the genome, and can be manifested or detected as differences in nucleic acid sequences, gene expression, including, for example transcription, processing, translation, transport, protein processing, trafficking, DNA synthesis, expressed proteins, other gene products or products of biochemical pathways or in post-translational modifications and any other differences manifested amongst members of a population. A single nucleotide polymorphism (SNP) refers to a polymorphism that arises as the result of a single base change, such as an insertion, deletion or change in a base.

A polymorphic marker or site is the locus at which divergence occurs. Such site may be as small as one base pair (an SNP). Polymorphic markers include, but are not limited to, restriction fragment length polymorphisms, variable number of tandem repeats (VNTR's), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats and other repeating patterns, simple sequence repeats and insertional elements, such as Alu. Polymorphic forms also are manifested as different mendelian

15

20

25

alleles for a gene. Polymorphisms may be observed by differences in proteins, protein modifications, RNA expression modification, DNA and RNA methylation, regulatory factors that alter gene expression and DNA replication, and any other manifestation of alterations in genomic nucleic acid or organelle nucleic acids.

As used herein, a healthy population, refers to a population of organisms, including but are not limited to, animals, bacteria, viruses, parasites, plants, eubacteria, and others, that are disease free. The concept of disease-free is a function of the selected organism. For example, for mammals it refers to a subject not manifesting any disease state. Practically a healthy subject, when human, is defined as human donor who passes blood bank criteria to donate blood for eventual use in the general population. These criteria are as follows: free of detectable viral, bacterial, mycoplasma, and parasitic infections; not anemic; and then further selected based upon a questionnaire regarding history (see Figure 3). Thus, a healthy population represents an unbiased population of sufficient health to donate blood according to blood bank criteria, and not further selected for any disease state. Typically such individuals are not taking any medications. For plants, for example, it is a plant population that does not manifest diseases pathology associated with plants. For bacteria it is a bacterial population replicating without environmental stress, such as selective agents, heat and other pathogens.

As used herein, a healthy database (or healthy patient database) refers to a database of profiles of subjects that have not been pre-selected for any particular disease. Hence, the subjects that serve as the source of data for the database are selected, according to predetermined criteria, to be healthy. In contrast to other such databases that have been pre-selected for subjects with a particular disease or other characteristic, the subjects for the database provided herein are not so-selected. Also, if the subjects do manifest a disease or other condition, any polymorphism discovered or characterized should be related to an independent disease or condition. In a preferred embodiment, where the subjects are human, a healthy subject manifests no disease symptoms and meets criteria, such as those set by blood banks for blood donors.

WO 01/27857 PCT/US00/28413

-13-

Thus, the subjects for the database are a population of any organism, including, but are not limited to, animals, plants, bacteria, viruses, parasites and any other organism or entity that has nucleic acid. Among preferred subjects are mammals, preferably, although not necessarily, humans. Such a database can capture the diversity of the a population, thus providing for discovery of rare polymorphisms.

As used herein, a profile refers to information relating to, but not limited to and not necessarily including all of, age, sex, ethnicity, disease history, family history, phenotypic characteristics, such as height and weight and other relevant parameters. A sample collect information form is shown in Figure 22, which illustrates profile intent.

As used herein, a disease state is a condition or abnormality or disorder that may be inherited or result from environmental stresses, such as toxins, bacterial, fungal and viral infections.

10

15

20

25

As used herein, set of non-selected subjects means that the subjects have not been pre-selected to share a common disease or other characteristic. They can be selected to be healthy as defined herein.

As used herein, a phenotype refers to a set of parameters that includes any distinguishable trait of an organism. A phenotype can be physical traits and can be, in instances in which the subject is an animal, a mental trait, such as emotional traits. Some phenotypes can be determined by observation elicited by questionnaires (see, e.g., Figures 3 and 22) or by referring to prior medical and other records. For purposes herein, a phenotype is a parameter around which the database can be sorted.

As used herein, a parameter is any input data that will serve as a basis for sorting the database. These parameters will include phenotypic traits, medical histories, family histories and any other such information elicited from a subject or observed about the subject. A parameter may describe the subject, some historical or current environmental or social influence experienced by the subject, or a condition or environmental influence on someone related to the subject. Parameters include, but are not limited to, any of those described herein, and known to those of skill in the art.

20

25

As used herein, haplotype refers referes to two or polymorphism located on a single DNA strand. Hence, haplotyping refers to identification of two or more polymorphisms on a single DNA strand. Haplotypes can be indicative of a phenotype. For some disorders a single polymorphism may suffice to indicate a trait; for others a plurality (i.e., a haplotype) may be needed. Haplotyping can be performed by isolating nucleic acid and separating the strands. In addition, when using enzymes such a certain nucleases, that produce, different size fragments from each strand, strand separation is not needed for haplotyping.

As used herein, used herein, pattern with reference to a mass spectrum or mass spectrometric analyses, refers to a characteristic distribution and number of signals (such peaks or digital representations thereof).

As used herein, signal in the context of a mass spectrum and analysis thereof refers to the output data, which the number or relative number of moleucles having a particular mass. Signals include "peaks" and digital representations thereof.

As used herein, adaptor, when used with reference to haplotyping use Fen ligase, refers to a nucleic acid that specifically hybridizes to a polymorphism of insterest. An adaptor can be partially double-stranded. An adaptor complex is formed when an adaptor hybridizes to its target.

As used herein, a target nucleic acid refers to any nucleic acid of interest in a sample. It can contain one or more nucleotides.

As used herein, standardless analysis refers to a determination based upon an internal standard. For example, the frequency of a polymorphism can be determined herein by comparing signals within a single mass spectrum.

As used herein, amplifying refers to means for increasing the amount of a bipolymer, especially nucleic acids. Based on the 5' and 3' primers that are chosen, amplication also serves to restrict and define the region of the genome which is subject to analysis. Amplification can be by any means known to those skilled in the art, including use of the polymerase chain reaction (PCR) etc.

Amplification, e.g., PCR must be done quantitatively when the frequency of polymorphism is required to be determined.

15

25

30

As used herein, cleaving refers to non-specific and specific fragmentation of a biopolymer.

As used herein, multiplexing refers to the simultaneous detection of more than one polymorphism. Methods for performing multiplexed reactions, particularly in conjunction with mass spectrometry are known (see, e.g., U.S. Patent Nos. 6,043,031, 5,547,835 and International PCT application No. WO 97/37041).

As used herein, reference to mass spectrometry encompasss any suitable mass spectrometric format known to those of skill in the art. Such formats cinlude, but are not limited to, Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI (see, e.g., published International PCT application No.99/57318 and U.S. Patent No. 5,118,937), Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof. MALDI, particular UV and IR, are among the preferred formats.

As used herein, mass spectrum refers to the presentation of data obtained from analyzing a biopolymer or fragment thereof by mass spectrometry either graphically or encoded numerically.

As used herein, a blood component is a component that is separated from blood and includes, but is not limited to red blood cells and platelets, blood clotting factors, plasma, enzymes, plasminogen, immunoglobulins. A cellular blood component is a component of blood, such as a red blood cell, that is a cell. A blood protein is a protein that is normally found in blood. Examples of such proteins are blood factors VII and VIII. Such proteins and components are well-known to those of skill in the art.

As used herein, plasma can be prepared by any method known to those of skill in the art. For example, it can be prepared by centrifuging blood at a force that pellets the red cells and forms an interface between the red cells and the buffy coat, which contains leukocytes, above which is the plasma. For example, typical platelet concentrates contain at least about 10% plasma.

Blood may be separated into its components, including, but not limited to, plasma, platelets and red blood cells by any method known to those of skill in the art. For example, blood can be centrifuged for a sufficient time and at a

15

20

25

30

sufficient acceleration to form a pellet containing the red blood cells. Leukocytes collect primarily at the interface of the pellet and supernatant in the buffy coat region. The supernatant, which contains plasma, platelets, and other blood components, may then be removed and centrifuged at a higher acceleration, whereby the platelets pellet.

As used herein, p53 is a cell cycle control protein that assesses DNA damage and acts as a transcription factor regulation gene which control cell growth, DNA repair and apoptosis. The p53 mutations have been found in a wide variety of different cancers, including all of the different types of leukemia, with varying frequency. The loss of normal p53 functions results in genomic instability and uncontrolled growth of the host cell.

As used herein, p21 is a cyclin-dependent kinase inhibitor, associated with G1 phase arrest of normal cells. Expression triggers apoptosis or programmed cell death and has been associated with Wilms' tumor, a pediatric kidney cancer.

As used herein, Factor VII is a serine protease involved the extrinsic blood coagulation cascade. This factor is activated by thrombin and works with tissue factor (Factor III) in the processing of Factor X to Factor Xa. Evidence has supported an association between polymorphisms in the gene and increase Factor VII activity which can result in an elevated risk of ischemic cardiovascular disease including myocardial infarction.

As used herein, a relational database stores information in a form representative of matrices, such as two-dimensional tables, including rows and columns of data, or higher dimensional matrices. For example, in one embodiment, the relational database has separate tables each with a parameter. The tables are linked with a record number, which also acts as an index. The database can be searched or sorted by using data in the tables and is stored in any suitable storage medium, such as floppy disk, CD rom disk, hard drive or other suitable medium.

As used herein, a bar codes refers any array of optically readable marks of any desired size and shape that are arranged in a reference context or frame of, preferably, although not necessarily, one or more columns and one or more

rows. For purposes herein, the bar code refers to any symbology, not necessary "bar" but may include dots, characters or any symbol or symbols.

As used herein, symbology refers to an identifier code or symbol, such as a bar code, that is linked to a sample. The index will reference each such symbology. The symbology is any code known or designed by the user. The symbols are associated with information stored in the database. For example, each sample can be uniquely identified with an encoded symbology. The parameters, such as the answers to the questions and subsequent genotypic and other information obtained upon analysis of the samples is included in the database and associated with the symbology. The database is stored on any suitable recording medium, such as a hard drive, a floppy disk, a tape, a CD ROM, a DVD disk and any other suitable medium.

DATABASES

15

20

25

30

Human genotyping is currently dependent on collaborations with hospitals, tissues banks and research institutions that provide samples of disease tissue. This approach is based on the concept that the onset and/or progression of diseases can be correlated with the presence of a polymorphisms or other genetic markers. This approach does not consider that disease correlated with the presence of specific markers and the absence of specific markers. It is shown herein that identification and scoring of the appearance and disappearance of markers is possible only if these markers are measured in the background of healthy subjects where the onset of disease does not mask the change in polymorphism occurrence. Databases of information from disease populations suffer from small sample size, selection bias and heterogeneity. The databases provided herein from healthy populations solve these problems by permitting large sample bands, simple selection methods and diluted heterogeneity.

Provided herein are first databases of parameters, associated with non-selected, particularly healthy, subjects. Also provided are combinations of the databases with indexed samples obtained from each of the subjects. Further provided are databases produced from the first databases. These contain in addition to the original parameters information, such as genotypic information,

20

25

30

including, but are not limited to, genomic sequence information, derived from the samples.

The databases, which are herein designated healthy databases, are so-designated because they are not obtained from subjects pre-selected for a particular disease. Hence, although individual members may have a disease, the collection of individuals is not selected to have a particular disease.

The subjects from whom the parameters are obtained comprise either a set of subjects who are randomly selected across, preferably, all populations, or are pre-selected to be disease-free or healthy. As a result, the database is not selected to be representative of any pre-selected phenotype, genotype, disease or other characteristic. Typically the number of subjects from which the database is prepared is selected to produce statistically significant results when used in the methods provided herein. Preferably, the number of subjects will be greater than 100, more preferably greater than 200, yet more preferably greater than 1000. The precise number can be empirically determined based upon the frequency of the parameter(s) that be used to sort the database. Generally the population can have at least 50, at least 100, at least 200, at least 500, at least 1000, at least 500 or at least 10,000 or more subjects.

Upon identification of a collection of subjects, information about each subject is recorded and associated with each subject as a database. The information associated with each of the subjects, includes, but is not limited to, information related to historical characteristics of the subjects, phenotypic characteristics and also genotypic characteristics, medical characteristics and any other traits and characteristics about the subject that can be determined. This information will serve as the basis for sorting the database.

In an exemplary embodiment, the subjects are mammals, such as humans, and the information relates to one or more of parameters, such as age, sex, medical history, ethnicity and any other factor. Such information, when the animals are humans, for example, can be obtained by a questionnaire, and by observations about the individual, such as hair color, eye color and other characteristics. Genotypic information will be obtained from tissue or other body and body fluid samples from the subject.

PCT/US00/28413

10

25

The healthy genomic database can include profiles and polymorphisms from healthy individuals from a library of blood samples where each sample in the library is an individual and separate blood or other tissue sample. Each sample in the database is profiled as to the sex, age, ethnic group, and disease history of the donor.

The databases are generated by first identifying healthy populations of subjects and obtaining information about each subject that will serve as the sorting parameters for the database. This information is preferably entered into a storage medium, such as the memory of a computer.

The information obtained about each subject in a population used for generating the database is stored in a computer memory or other suitable storage medium. The information is linked to an identifier associated with each subject. Hence the database will identify a subject, for example by a datapoint representative of a bar code, and then all information, such as the information from a questionnaire, regarding the individual is associated with the datapoint. As the information is collected the database is generated.

Thus, for example, profile information, such as subject histories obtained from questionnaires, is collected in the database. The resulting database can be sorted as desired, using standard software, such as by age, sex and/or ethnicity. An exemplary questionnaire for subjects from whom samples are to be obtained is shown in Figures 22A-D. Each questionnaire preferably is identified by a bar code, particularly a machine readable bar code for entry into the database. After a subject provides data and is deemed to be healthy (i.e., meets standards for blood donation), the data in the questionnaire is entered into the database and is associated with the bar code. A tissue, cell or blood sample is obtained from the subject.

Figure 4 exemplifies processing and tracking of blood sample components. Each component is tracked with a bar code, dated, is entered into the database and associated with the subject and the profile of the subject. Typically, the whole blood is centrifuged to produce plasma, red blood cells (which pellet) and leukocytes found in the buffy coat which layers in between.

25

30

Various samples are obtained and coded with a bar code and stored for use as needed.

Samples are collected from the subjects. The samples include, but are not limited to, tissues, cells, and fluids, such as nucleic acid, blood, plasma, amniotic fluid, synovial fluid, urine, saliva, aqueous humor, sweat, sperm samples and cerebral spinal fluid. It is understood that the particular set of samples depends upon the organisms in the population.

Once samples are obtained the collection can be stored and, in preferred embodiments, each sample is indexed with an identifier, particularly a machine readable code, such as a bar code. For analyses, the samples or components of the samples, particularly biopolymers and small molecules, such as nucleic acids and/or proteins and metabolites, are isolated.

After samples are analyzed, this information is entered into the database in the memory of the storage medium and associated with each subject. This information includes, but is not limited to, genotypic information. Particularly, nucleic acid sequence information and other information indicative of polymorphisms, such as masses of PCR fragments, peptide fragment sequences or masses, spectra of biopolymers and small molecules and other indicia of the structure or function of a gene, gene product or other marker from which the existence of a polymorphism within the population can be inferred.

In an exemplary embodiment, a database can be derived from a collection of blood samples. For example, Figure 1 (see, also Figure 10) shows the status of a collection of over 5000 individual samples. The samples were processed in the laboratory following SOP (standard operating procedure) guidelines. Any standard blood processing protocol may be used.

For the exemplary database described herein, the following criteria were used to select subjects:

No testing is done for infectious agents.

Age: At least 17 years old

Weight: Minimum of 110 pounds

Permanently Disqualified:

History of hepatitis (after age 11)

Leukemia Lymphoma Human immunodeficiency virus (HIV),AIDS Chronic kidney disease

Temporarily Disqualified:

Pregnancy - until six weeks after delivery, miscarriage or abortion
Major surgery or transfusions - for one year
Mononucleosis - until complete recovery
Prior whole blood donation - for eight weeks
Antibiotics by injection for one week; by mouth, for forty-eight hours,
except antibiotics for skin complexion;

5 year Deferment:

Internal cancer and skin cancer if it has been removed, is healed and there is no recurrence

These correspond to blood bank criteria for donating blood and represent a healthy population as defined herein for a human healthy database.

Structure of the database

Any suitable database structure and format known to those of skill in the art may be employed. For example, a relational database is a preferred format in which data is stored as matrices or tables of the parameters linked by an indexer that identifies each subject. Software for preparing and manipulating, including sorting the database, can be readily developed or adapted from commercially available software, such as Microsoft Access.

Quality control

Quality control procedures can be implemented. For example, after

collection of samples, the quality of the collection in the bank can be assessed. For example, mix-up of samples can be checked by testing for known markers, such as sex. After samples are separated by ethnicity, samples are randomly tested for a marker associated with a particular ethnicity, such as HLA DQA1 group specific component, to assess whether the samples have been properly sorted by ethnic group. An exemplary sample bank is depicted in Figure 4.

15

20

25

30

Obtaining genotypic data and other parameters for the database

After informational and historical parameters are entered into the database, material from samples obtained from each subject, is analyzed. Analyzed material include proteins, metabolites, nucleic acids, lipids and any other desired constituent of the material. For example, nucleic acids, such as genomic DNA, can be analyzed by sequencing.

Sequencing can be performed using any method known to those of skill in the art. For example, if a polymorphism is identified or known, and it is desired to assess its frequency or presence among the subjects in the database, the region of interest from each sample can be isolated, such as by PCR or restriction fragments, hybridization or other suitable method known to those of skill in the art and sequenced. For purposes herein, sequencing analysis is preferably effected using mass spectrometry (see, e.g., U.S. Patent Nos. 5,547,835, 5,622,824, 5,851,765, and 5,928,906). Nucleic acids can also be sequence by hybridization (see, e.g., U.S. Patent Nos. 5,503,980, 5,631,134, 5,795,714) and including analysis by mass spectrometry (see, U.S. application Serial Nos. 08/419,994 and 09/395,409).

In other detection methods, it is necessary to first amplify prior to identifying the allelic variant. Amplification can be performed, e.g., by PCR and/or LCR, according to methods known in the art. In one embodiment, genomic DNA of a cell is exposed to two PCR primers and amplification for a number of cycles sufficient to produce the required amount of amplified DNA. In preferred embodiments, the primers are located between 150 and 350 base pairs apart.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J. C. et al., 1990, Proc. Natl. Acad. Sci. U.S.A. 87:1874-1878), transcriptional amplification system (Kwoh, D. Y. et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:1173-1177), Q-Beta Replicase (Lizardi, P. M. et al., 1988, Bio/Technology 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially

30

useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

Nucleic acids can also be analyzed by detection methods and protocols, particularly those that rely on mass spectrometry (see, e.g., U.S. Patent No. 5 5,605,798, 6,043,031, allowed copending U.S. application Serial No. 08/744,481, U.S. application Serial No. 08/990,851 and International PCT application No. WO 99/31273, International PCT application No. WO 98/20019). These methods can be automated (see, e.g., copending U.S. application Serial No. 09/285,481 and published International PCT application No. PCT/US00/08111, which describes an automated process line). Preferred among the methods of analysis herein are those involving the primer oligo base extension (PROBE) reaction with mass spectrometry for detection (described herein and elsewhere, see e.g., U.S. Patent No. 6,043,031; see, also U.S. application Serial Nos. 09/287,681, 09/287,682, 09/287,141 and 09/287,679, allowed copending U.S. application Serial No. 08/744,481, International PCT application No. PCT/US97/20444, published as International PCT application No. WO 98/20019, and based upon U.S. application Serial Nos. 08/744,481, 08/744,590, 08/746,036, 08/746,055, 08/786,988, 08/787,639, 08/933,792, 08/746,055, 08/786,988 and 08/787,639; see, also U.S. application Serial No. 09/074,936, U.S. Patent No. 6,024,925, and U.S. application Serial Nos. 08/746,055 and 08/786,988, and published International PCT application No. WO 98/20020)

A preferred format for performing the analyses is a chip based format in which the biopolymer is linked to a solid support, such as a silicon or silicon-coated substrate, preferably in the form of an array. More preferably, when analyses are performed using mass spectrometry, particularly MALDI, small nanoliter volumes of sample are loaded on, such that the resulting spot is about, or smaller than, the size of the laser spot. It has been found that when this is achieved, the results from the mass spectrometric analysis are quantitative. The area under the signals in the resulting mass spectra are proportional to concentration (when normalized and corrected for background). Methods for preparing and using such chips are described in U.S. Patent No. 6,024,925, co-

15

20

25

pending U.S. application Serial Nos. 08/786,988, 09/364,774, 09/371,150 and 09/297,575; see, also U.S. application Serial No. PCT/US97/20195, which published as WO 98/20020. Chips and kits for performing these analyses are commercially available from SEQUENOM under the trademark MassARRAY. MassArray relies on the fidelity of the enzymatic primer extension reactions combined with the miniaturized array and MALDI-TOF (Matrix-Assisted Laser Desorption Ionization-Time of Flight) mass spectrometry to deliver results rapidly. It accurately distinguishes single base changes in the size of DNA fragments

The methods provided herein permit quantitative determination of alleles. The areas under the signals in the mass spectra can be used for quantitative determinations. The frequency is determined from the ratio of the signal to the total area of all of the spectrum and corrected for background. This is possible because of the PROBE technology as described in the above applications incorporated by reference herein.

associated with genetic variants without tags.

Additional methods of analyzing nucleic acids include amplification- based methods including polymerase chain reaction (PCR), ligase chain reaction (LCR), mini-PCR, rolling circle amplification, autocatalytic methods, such as those using QB replicase, TAS, 3SR, and any other suitable method known to those of skill in the art.

Other methods for analysis and identification and detection of polymorphisms, include but are not limited to, allele specific probes, Southern analyses, and other such analyses. .:

The methods described below provide ways to fragment given amplified or non-amplified nucleotide sequences thereby producing a set of mass signals when mass spectrometry is used to analyze the fragment mixtures. Amplified fragments are yielded by standard polymerase chain methods (US 4,683,195 and 4,683,202). The fragmentation method involves the use of enzymes that cleave single or double strands of DNA and enzymes that ligate 30 DNA. The cleavage enzymes can be glycosylases, nickases, and site-specific and non site-specific nucleases with the most preferred enzymes being glycosylases, nickases, and site-specific nucleases.

PCT/US00/28413

-25-

Glycosylase Fragmentation Method

10

15

20

25

DNA glycosylases specifically remove a certain type of nucleobase from a given DNA fragment. These enzymes can thereby produce abasic sites, which can be recognized either by another cleavage enzyme, cleaving the exposed phosphate backbone specifically at the abasic site and producing a set of nucleobase specific fragments indicative of the sequence, or by chemical means, such as alkaline solutions and or heat. The use of one combination of a DNA glycosylase and its targeted nucleotide would be sufficient to generate a base specific signature pattern of any given target region.

Numerous DNA glcosylases are known, For example, a DNA glycosylase can be uracil-DNA glycolsylase (UDG), 3-methyladenine DNA glycosylase, 3-methyladenine DNA glycosylase II, pyrimidine hydrate-DNA glycosylase, FaPy-DNA glycosylase, thymine mismatch-DNA glycosylase, hypoxanthine-DNA glycosylase, 5-Hydroxymethyluracil DNA glycosylase (HmUDG), 5-Hydroxymethylcytosine DNA glycosylase, or 1,N6-ethenoadenine DNA glycosylase (see, e.g.,, U.S. Patent Nos. 5,536,649, 5,888, 795, 5,952,176 and 6,099,553, International PCT application Nos. WO 97/03210, WO 99/54501; see, also, Eftedal et al. (1993) Nucleic Acids Res 21:2095-2101, Bjelland and Seeberg (1987) Nucleic Acids Res. 15:2787-2801, Saparbaev et al. (1995) Nucleic Acids Res. 23:3750-3755, Bessho (1999) Nucleic Acids Res. 27:979-983) corresponding to the enzyme's modified nucleotide or nucleotide analog target. A preferred glycosylase is uracil-DNA glycolsylase (UDG).

Uracil, for example, can be incorporated into an amplified DNA molecule by amplifying the DNA in the presence of normal DNA precursor nucleotides (e.g. dCTP, dATP, and dGTP) and dUTP. When the amplified product is treated with UDG, uracil residues are cleaved. Subsequent chemical treatment of the products from the UDG reaction results in the cleavage of the phosphate backbone and the generation of nucleobase specific fragments. Moreover, the separation of the complementary strands of the amplified product prior to glycosylase treatment allows complementary patterns of fragmentation to be generated. Thus, the use of dUTP and Uracil DNA glycosylase allows the generation of T specific fragments for the complementary strands, thus providing

20

25

30

information on the T as well as the A positions within a given sequence. Similar to this, a C-specific reaction on both (complementary) strands (i.e. with a C-specific glycosylase) yields information on C as well as G positions within a given sequence if the fragmentation patterns of both amplification strands are analyzed separately. Thus, with the glycosylase method and mass spectrometry, a full series of A, C, G and T specific fragmentation patterns can be analyzed.

Nickase Fragmentation Method

A DNA nickase, or DNase, can be used recognize and cleave one strand of a DNA duplex. Numerous nickases are known. Among these, for example, are nickase NY2A nickase and NYS1 nickase (Megabase) with the following cleavage sites:

NY2A: 5'...R AG...3'

3'...Y TC...5' where R = A or G and Y = C or T

NYS1: 5'... CC[A/G/T]...3'

3'... GGIT/C/A1...5'.

Fen-Ligase Fragmentation Method

The Fen-ligase method involves two enzymes: Fen-1 enzyme and a ligase. The Fen-1 enzyme is a site-specific nuclease known as a "flap" endonuclease (US 5,843,669, 5,874,283, and 6,090,606). This enzymes recognizes and cleaves DNA "flaps" created by the overlap of two oligonucleotides hybridized to a target DNA strand. This cleavage is highly specific and can recognize single base pair mutations, permitting detection of a single homologue from an individual heterozygous at one SNP of interest and then genotyping that homologue at other SNPs occurring within the fragment. Fen-1 enzymes can be Fen-1 like nucleases e.g. human, murine, and Xenopus XPG enzymes and yeast RAD2 nucleases or Fen-1 endonucleases from, for example, *M. jannaschii*, *P. furiosus*, and *P. woesei*. Among preferred enzymes are the Fen-1 enzymes.

The ligase enzyme forms a phosphodiester bond between two double stranded nucleic acid fragments. The ligase can be DNA Ligase I or DNA Ligase III (see, e.g., U.S. Patent Nos. US 5,506,137, 5,700,672, 5,858,705 and 5,976,806; see, also, Waga, et al. (1994) J. Biol. Chem. 269:10923-10934, Li

5 .

10

20

25

et al. (1994) Nucleic Acids Res. 22:632-638, Arrand et al. (1986) J. Biol. Chem. 261:9079-9082, Lehman (1974) Science 186:790-797, Higgins and Cozzarelli (1979) Methods Enzymol. 68:50-71, Lasko et al. (1990) Mutation Res. 236:277-287, and Lindahl and Barnes (1992) Ann. Rev. Biochem. 61:251-281).

Thermostable ligase (Epicenter Technologies), where "thermostable" denotes that the ligase retains activity even after exposure to temperatures necessary to separate two strands of DNA, are among preferred ligases for use herein.

Type IIS Enzyme Fragmentation Method

Restriction enzymes bind specifically to and cleave double-stranded DNA at specific sites within or adjacent to a particular recognition sequence. These enzymes have been classified into three groups (e.g. Types I, II, and III) as known to those of skill in the art. Because of the properties of type I and type III enzymes, they have not been widely used in molecular biological applications. Thus, for this invention type II enzymes are preferred. Of the thousands of restriction enzymes known in the arts, there are 179 different type II specificities. Of the 179 unique type II restriction endonucleases, 31 have a 4-base recognition sequence, 11 have a 5-base recognition sequence, 127 have a 6-base recognition sequence, and 10 have recognition sequences of greater than six bases (US 5,604,098). Of category type II enzymes, type IIS is preferred.

Type IIS enzymes can be Alw XI, Bbv I, Bce 83, Bpm I, Bsg I, Bsm AI, Bsm FI, Bsa I, Bcc I, Bcg I, Ear I, Eco 571, Esp 31, Fau I, Fok I, Gsu I, Hga I, Mme I, Mbo II, Sap I, and the like. The preferred type IIS enzyme is Fok I.

The Fok I enzyme endonuclease is an exemplary well characterized member of the Type IIS class (see, e.g., U.S. Patent Nos. 5,714,330, 5,604,098, 5,436,150, 6,054,276 and 5,871,911; see, also, Szybalski et al. (1991) Gene 100:13-26, Wilson and Murray (1991) Ann. Rev. Genet. 25:585-627, Sugisaki et al. (1981) Gene 16:73-78, Podhajska and Szalski (1985) Gene 40:175-182. Fok I recognizes the sequence 5'GGATG-3' and cleaves DNA accordingly. Type IIS restriction sites can be introduced into DNA targets by incorporating the site into primers used to amplify such targets. Fragments produced by digestion with Fok I are site specific and can be analyzed by mass

15

20

spectrometry methods such as MALDI-TOF mass spectrometry, ESI-TOF mass spectrometry, and any other type of mass spectrometry well known to those of skill in the art.

Once a polymorphism has been found to correlatate with a parameter such as age. The possibility of false results due to allelic dropout is examined by doing comparative PCR in an adjacent region of the genome.

Analyses

In using the database, allelic frequencies can be determined across the population by analyzing each sample in the population individually, determining the presence or absence of allele or marker of interest in each individual sample, and then determining the frequency of the marker in the population. The database can then be sorted (stratified) to identify any correlations between the allele and a selected parameter using standard statistical analysis. If a correlation is observed, such as a decrease in a particular marker with age or correlation with sex or other parameter, then the marker is a candidate for further study, such as genetic mapping to identify a gene or pathway in which it is involved. The marker may then be correlated, for example, with a disease. Haplotying can also be carried out. Genetic mapping can be effected using standard methods and may also require use of databases of others, such as databases previously determined to be associated with a disorder.

Exemplary analyses have been performed and these are shown in the figures, and discussed herein.

Sample pooling

It has been found that using the databases provided herein, or any other database of such information, substantially the same frequencies that were obtained by examining each sample separately can be obtained by pooling samples, such as in batches of 10, 20, 50, 100, 200, 500, 1000 or any other number. A precise number may be determined empirically if necessary, and can be as low as 3.

15

20

25

30

In one embodiment, the frequency of genotypic and other markers can be obtained by pooling samples. To do this a target population and a genetic variation to be assessed is selected, a plurality of samples of biopolymers are obtained from members of the population, and the biopolymer from which the marker or genotype can be inferred is determined or detected. A comparison of samples tested in pools and individually and the sorted results therefrom are shown in Figure 9, which shows frequency of the factor VII Allele 353Q. Figure 10 depicts the frequency of the CETP Allele CETP in pooled versus individual samples. Figure 15 shows ethnic diversity among various ethnic groups in the database using pooled DNA samples to obtain the data. Figures 12-14 show mass spectra for these samples.

Pooling of test samples has application not only to the healthy databases provided herein, but also to use in gathering data for entry into any database of subjects and genotypic information, including typical databases derived from diseased populations. What is demonstrated herein, is the finding that the results achieved are statistically the same as the results that would be achieved if each sample is analyzed separately. Analysis of pooled samples by a method, such as the mass spectrometric methods provided herein, permits resolution of such data and quantitation of the results.

For factor VII the R53Q acid polymorphism was assessed. In Figure 9, the "individual" data represent allelic frequency observed in 92 individuals reactions. The pooled data represent the allelic frequency of the same 92 individuals pooled into a single probe reaction. The concentration of DNA in the samples of individual donors is 250 nanograms. The total concentration of DNA in the pooled samples is also 250 nanograms, where the concentration of any individual DNA is 2.7 nanograms.

It also was shown that it is possible to reduce the DNA concentration of individuals in a pooled samples from 2.7 nanograms to 0.27 nanograms without any change in the quality of the spectrum or the ability to quantitate the amount of sample detected. Hence low concentrations of sample may be used in the pooling methods.

25

30

Use of the databases and markers identified thereby

The successful use of genomics requires a scientific hypothesis (i.e., common genetic variation, such as a SNP), a study design (i.e., complex disorders), samples and technology, such as the chip-based mass spectrometric analyses (see, e.g., U.S. Patent No. 5,605,798, U.S. Patent No. 5,777,324, U.S. Patent No. 6,043,031, allowed copending U.S. application Serial No. 08/744,481, U.S. application Serial No. 08/990,851, International PCT application No. WO 98/20019, copending U.S. application Serial No. 09/285,481, which describes an automated process line for analyses; see, also, U.S. application Serial Nos. 08/617,256, 09/287,681, 09/287,682, 09/287,141 10 and 09/287,679, allowed copending U.S. application Serial No. 08/744,481, International PCT application No. PCT/US97/20444, published as International PCT application No. WO 98/20019, and based upon U.S. application Serial Nos. 08/744,481, 08/744,590, 08/746,036, 08/746,055, 08/786,988, 08/787,639, 08/933,792, 08/746,055, 09/266,409, 08/786,988 and 08/787,639; see, also U.S. application Serial No. 09/074,936). All of these aspects can be used in conjunction with the databases provided herein and samples in the collection.

The databases and markers identified thereby can be used, for example, for identification of previously unidentified or unknown genetic markers and to identify new uses for known markers. As markers are identified, these may be entered into the database to use as sorting parameters from which additional correlations may be determined.

Previously unidentified or unknown genetic markers

The samples in the healthy databases can be used to identify new polymorphisms and genetic markers, using any mapping, sequencing, amplification and other methodologies, and in looking for polymorphisms among the population in the database. The thus-identified polymorphism can then be entered into the database for each sample, and the database sorted (stratified) using that polymorphism as a sorting parameter to identify any patterns and correlations that emerge, such as age correlated changes in the frequency of the identified marker. If a correlation is identified, the locus of the marker can be mapped and its function or effect assessed or deduced.

15

20

25

30

Thus, the databases here provide means for:

identification of significantly different allelic frequencies of genetic factors by comparing the occurrence or disappearance of the markers with increasing age in population and then associating the markers with a disease or a biochemical pathway;

identification of significantly different allelic frequencies of disease causing genetic factors by comparing the male with the female population or comparing other selected stratified populations and associating the markers with a disease or a biochemical pathway;

identification of significantly different allelic frequencies of disease causing genetic factors by comparing different ethnic groups and associating the markers with a disease or a biochemical pathway that is known to occur in high frequency in the ethnic group;

profiling potentially functional variants of genes through the general panmixed population stratified according to age, sex, and ethnic origin and thereby demonstrating the contribution of the variant genes to the physical condition of the investigated population;

identification of functionally relevant gene variants by gene disequilibrium analysis performed within the general panmixed population stratified according to age, sex, and ethnic origin and thereby demonstrating their contribution to the physical condition of investigated population;

identification of potentially functional variants of chromosomes or parts of chromosomes by linkage disequilibrium analysis performed within the general panmixed population stratified according to age, sex, and ethnic origin and thereby demonstrating their contribution to the physical condition of investigated population.

Uses of the identified markers and known markers

The databases may also be used in conjunction with known markers and sorted to identify any correlations. For example, the databases can be used for:

determination and evaluation of the penetrance of medically relevant polymorphic markers;

15

20

25

30

determination and evaluation of the diagnostic specificity of medically relevant genetic factors;

determination and evaluation of the positive predictive value of medically relevant genetic factors;

determination and evaluation of the onset of complex diseases, such as, but are not limited to, diabetes, hypertension, autoimmune diseases, arteriosclerosis, cancer and other diseases within the general population with respect to their causative genetic factors;

delineation of the appropriate strategies for preventive disease treatment; delineation of appropriate timelines for primary disease intervention; validation of medically relevant genetic factors identified in isolated populations regarding their general applicability;

validation of disease pathways including all potential target structures identified in isolated populations regarding their general applicability; and

validation of appropriate drug targets identified in isolated populations regarding their general applicability.

Among the diseases and disorders for which polymorphisms may be linked include, those linked to inborn errors of metabolism, acquired metabolic disorders, intermediary metabolism, oncogenesis pathways, blood clotting pathways, and DNA synthetic and repair pathways DNA repair/replication/transcription factors and activities, e.g., such as genes related to oncogenesis, aging and genes involved in blood clotting and the related biochemical pathways that are related to thrombosis, embolism, stroke, myocardial infarction, angiogenesis and oncogenesis.

For example, a number of diseases are caused by or involve deficient or defective enzymes in intermediary metabolism (see, e.g., Tables 1 and 2, below) that result, upon ingestion of the enzyme substrates, in accumulation of harmful metabolites that damage organs and tissues, particularly an infant's developing brain and other organs, resulting in mental retardation and other developmental disorders.

20

25

30

Identification of markers and genes for such disorders is of great interest.

Model systems

Several gene systems, p21, p53 and Lipoprotein Lipase polymorphism (N291S), were selected. The p53 gene is a tumor suppressor gene that is mutated in diverse tumor types. One common allelic variant occurs at codon 72. A polymorphism that has been identified in the p53 gene, i.e., the R72P allele, results in an amino acid exchange, arginine to proline, at codon 72 of the gene.

Using diseased populations, it has been shown that there are ethnic differences in the allelic distribution of these alleles among African-Americans and Caucasians in the U.S. The results here support this finding and also demonstrate that the results obtained with a healthy database are meaningful (see, Figure 7B).

The 291S allele leads to reduced levels of high density lipoprotein cholesterol (HDL-C) that is associated with an increased risk of males for arteriosclerosis and in particular myocardial infarction (see, Reymer *et al.* (1995) *Nature Genetics 10*:28-34).

Both genetic polymorphisms were profiled within a part of the Caucasian population-based sample bank. For the polymorphism located in the lipoprotein lipase gene a total of 1025 unselected individuals (436 males and 589 females) were tested. Genomic DNA was isolated from blood samples obtained from the individuals.

As shown in the Examples and figures, an exemplary database containing about 5000 subjects, answers to the questionnaire (see Figure 3), and genotypic information has been stratified. A particular known allele has been selected, and the samples tested for the marker using mass spectrometric analyses, particularly PROBE (see the EXAMPLES) to identify polymorphisms in each sample. The population in the database has been sorted according to various parameters and correlations have been observed. For example, FIGURES 2A-C, show sorting of the data by age and sex for the Lipoprotein Lipase gene in the Caucasian population in the database. The results show a decrease in the frequency of the allele with age in males but no such decrease in females. Other

25

30

alleles that have been tested against the database, include, alleles of p53, p21 and factor VII. Results when sorted by age are shown in the figures.

These examples demonstrate an effect of altered frequency of disease causing genetic factors within the general population. The scientific

5 interpretation of those results allows prediction of medical relevance of polymorphic genetic alterations. In addition, conclusions can be drawn with regard to their penetrance, diagnostic specificity, positive predictive value, onset of disease, most appropriate onset of preventive strategies, and the general applicability of genetic alterations identified in isolated populations to panmixed populations.

Therefore, an age- and sex-stratified population-based sample bank that is ethnically homogenous is a suitable tool for rapid identification and validation of genetic factors regarding their potential medical utility.

Exemplary computer system for creating, storing and processing the databases

Systems

Systems, including computers, containing the databases are provided herein. The computers and databases can be used in conjunction, for example, with the APL system (see, copending U.S. application Serial No. 09/285,481), which is an automated system for analyzing biopolymers, particularly nucleic acids. Results from the APL system can be entered into the database.

Any suitable computer system may be used. The computer system may be integrated into systems for sample analysis, such as the automated process line described herein (see, e.g., copending U.S. application Serial No. 09/285,481).

Figure 17 is a block diagram of a computer constructed in to provide and process the databases described herein. The processing that maintains the database and performs the methods and procedures may be performed on multiple computers all having a similar construction, or may be performed by a single, integrated computer. For example, the computer through which data is added to the database may be separate from the computer through which the database is sorted, or may be integrated with it. In either arrangement, the

20

25

computers performing the processing may have a construction as illustrated in Figure 17.

Figure 17 is a block diagram of an exemplary computer 1700 that maintains the database described above and performs the methods and procedures. Each computer 1700 operates under control of a central processor unit (CPU) 1702, such as a "Pentium" microprocessor and associated integrated circuit chips, available from Intel Corporation of Santa Clara, California, USA. A computer user can input commands and data from a keyboard and display mouse 1704 and can view inputs and computer output at a display 1706. The display is typically a video monitor or flat panel display device. The computer 1700 also includes a direct access storage device (DASD) 1707, such as a fixed hard disk drive. The memory 1708 typically comprises volatile semiconductor random access memory (RAM). Each computer preferably includes a program product reader 1710 that accepts a program product storage device 1712, from which the program product reader can read data (and to which it can optionally write data). The program product reader can comprise, for example, a disk drive, and the program product storage device can comprise removable storage media such as a magnetic floppy disk, an optical CD-ROM disc, a CD-R disc, a CD-RW disc, or a DVD data disc. If desired, the computers can be connected so they can communicate with each other, and with other connected computers, over a network 1713. Each computer 1700 can communicate with the other connected computers over the network 1713 through a network interface 1714 that enables communication over a connection 1716 between the network and the computer.

The computer 1700 operates under control of programming steps that are temporarily stored in the memory 1708 in accordance with conventional computer construction. When the programming steps are executed by the CPU 1702, the pertinent system components perform their respective functions. Thus, the programming steps implement the functionality of the system as described above. The programming steps can be received from the DASD 1707, through the program product reader 1712, or through the network connection 1716. The storage drive 1710 can receive a program product, read

20

30

programming steps recorded thereon and transfer the programming steps into the memory 1708 for execution by the CPU 1702. As noted above, the program product storage device 1710 can comprise any one of multiple removable media having recorded computer-readable instructions, including magnetic floppy disks and CD-ROM storage discs. Other suitable program product storage devices can include magnetic tape and semiconductor memory chips. In this way, the processing steps necessary for operation can be embodied on a program product.

Alternatively, the program steps can be received into the operating memory 1708 over the network 1713. In the network method, the computer receives data including program steps into the memory 1708 through the network interface 1714 after network communication has been established over the network connection 1716 by well-known methods that will be understood by those skilled in the art without further explanation. The program steps are then executed by the CPU 1702 to implement the processing of the Garment Database system.

It should be understood that all of the computers of the system preferably have a construction similar to that shown in Figure 17, so that details described with respect to the Figure 17 computer 1700 will be understood to apply to all computers of the system 1700. This is indicated by multiple computers 1700 shown connected to the network 1713. Any one of the computers 1700 can have an alternative construction, so long as they can communicate with the other computers and support the functionality described herein.

Figure 18 is a flow diagram that illustrates the processing steps performed using the computer illustrated in Figure 17, to maintain and provide access to the databases, such as for identifying polymorphic genetic markers. In particular, the information contained in the database is stored in computers having a construction similar to that illustrated in Figure 17. The first step for maintaining the database, as indicated in Figure 18, is to identify healthy members of a population. As noted above, the population members are subjects that are selected only on the basis of being healthy, and where the subjects are mammals, such as humans, they are preferably selected based upon apparent

20

25

30

health and the absence of detectable infections. The step of identifying is represented by the flow diagram box numbered 1802.

The next step, represented by the flow diagram box numbered 1804, is to obtain identifying and historical information and data relating to the identified members of the population. The information and data comprise parameters for each of the population members, such as member age, ethnicity, sex, medical history, and ultimately genotypic information. Initially, the parameter information is obtained from a questionnaire answered by each member, from whom a body tissue or body fluid sample also is obtained. The step of entering and storing these parameters into the database of the computer is represented by the flow diagram box numbered 1806. As additional information about each population member and corresponding sample is obtained, this information can be inputted into the database and can serve as a sorting parameter.

In the next step, represented by the flow diagram box numbered 1808, the parameters of the members are associated with an indexer. This step may be executed as part of the database storage operation, such as when a new data record is stored according to the relational database structure and is automatically linked with other records according to that structure. The step 1806 also may be executed as part of a conventional data sorting or retrieval process, in which the database entries are searched according to an input search or indexing key value to determine attributes of the data. For example, such search and sort techniques may be used to follow the occurrence of known genetic markers and then determine if there is a correlation with diseases for which they have been implicated. Examples of this use are for assessing the frequencies of the p53 and Lipoprotein Lipase polymorphisms.

Such searching of the database also may be valuable for identifying one or more genetic markers whose frequency changes within the population as a function of age, ethnic group, sex, or some other criteria. This can allow the identification of previously unknown polymorphisms and, ultimately, identification of a gene or pathway involved in the onset and progression of disease.

20

25

30

In addition, the database can be used for taking an identified polymorphism and ascertaining whether it changes in frequency when the data is sorted according to a selected parameter.

In this way, the databases and methods provided herein permit, among other things, identification of components, particularly key components, of a disease process by understanding its genetic underpinnings, and also an understanding of processes, such as individual drug responses. The databases and methods provided herein also can be used in methods involving elucidation of pathological pathways, in developing new diagnostic assays, identifying new potential drug targets, and in identifying new drug candidates.

Morbidity and/or early mortality associated polymorphisms

A database containing information provided by a population of healthy blood donors who were not selected for any particular disease to can be used to identify polymorphisms and the alleles in which they are present, whose frequency decreases with age. These may represent morbidity susceptibility markers and genes.

Polymorphisms of the genome can lead to altered gene function, protein function or genome instability. To identify those polymorphisms which have a clinical relevance/utility is the goal of a world-wide scientific effort. It can be expected that the discovery of such polymorphisms will have a fundamental impact on the identification and development of novel drug compounds to cure diseases. However, the strategy to identify valuable polymorphisms is cumbersome and dependent upon the availability of many large patient and control cohorts to show disease association. In particular, genes that cause a general risk of the population to suffer from any disease (morbidity susceptibility genes) will escape these case/control studies entirely.

Here described is a screening strategy to identify morbidity susceptibility-genes underlying a variety of different diseases. The definition of a morbidity susceptibility gene is a gene that is expressed in many different cell types or tissues (housekeeping gene) and its altered function can facilitate the expression of a clinical phenotype caused by disease-specific susceptibility genes that are involved in a pathway specific for this disorder. In other words, these morbidity

15

20

25

30

susceptibility genes predispose people to develop a distinct disease according to their genetic make-up for this disease.

Candidates for morbidity susceptibility genes can be found at the bottom level of pathways involving transcription, translation, heat-shock proteins, protein trafficking, DNA repair, assembly systems for subcellular structures (e.g. mitochondria, peroxysomes and other cellular microbodies), receptor signaling cascades, immunology, etc. Those pathways control the quality of life at the cellular level as well as for the entire organism. Mutations/polymorphisms located in genes encoding proteins for those pathways can reduce the fitness of cells and make the organism more susceptible to express the clinical phenotype caused by the action of a disease-specific susceptibility gene. Therefore, these morbidity susceptibility genes can be potentially involved in a whole variety of different complex diseases if not in all. Disease-specific susceptibility genes are involved in pathways that can be considered as disease-specific pathways like glucose-, lipid, hormone metabolism, etc.

The exemplified method permit, among other things, identification of genes and/or gene products involved in a man's general susceptibility to morbidity and/or mortality; use of these genes and/or gene products in studies to elucidate the genetic underpinnings of human diseases; use of these genes and/or gene products in combinatorial statistical analyses without or together with disease-specific susceptibility genes; use of these genes and/or gene products to predict penetrance of disease susceptibility genes; use of these genes and/or gene products in predisposition and/or acute medical diagnostics and use of these genes and/or gene products to develop drugs to cure diseases and/or to extend the life span of humans.

SCREENING PROCESS

The healthy population stratified by age, gender and ethnicity, etc. is a very efficient and a universal screening tool for morbidity associated genes. Changes of allelic frequencies in the young compared to the old population are expected to indicate putative morbidity susceptibility genes. Individual samples of this healthy population base can be pooled to further increase the throughput. In a proof of principle experiment pools of young and old Caucasian females and

males were applied to screen more than 400 randomly chosen single nucleotide polymorphisms located in many different genes. Candidate polymorphisms were identified if the allelic difference was greater than 8% between young and old for both or only one of the genders. The initial results were assayed again in at least one independent subsequent experiments. Repeated experiments are necessary to recognize unstable biochemical reactions, which occur with a frequency of about 2-3% and can mimic age-related allelic frequency differences. Average frequency differences and standard deviations are calculated after successful reproducibility of initial results. The final allelic frequency is then compared to a reference population of Caucasian CEPH sample pool. The result should show similar allelic frequencies in the young Caucasian population. Subsequently, the exact allele frequencies of candidates including genotype information were obtained by analyzing all individual samples. This procedure is straight forward with regard to time and cost. It enables the screening of an enormous number of SNPs. So far, several markers with a highly significant association to age were identified and described below.

In general at least 5 individual in a stratified population need to be screened to produce statistically significant results. The frequency of the allele is determined for an age stratified population. Chi square analysis is then performed on the allelic frequencies to determine if the difference between age groups is statistically significant. A p value less than of 0.1 is considered to represent a statistically significant difference. More preferably the p value should be less than 0.05.

Clinical Trials

20

25

30

The identification of markers whose frequency in a population decreases with age also allows for better designed and balanced clinical trials. Currently, if a clinical trial utilizes a marker as a significant endpoint in a study and the marker disappears with age, then the results of the study may be inaccurate. By using methods provided herein, it can be ascertained that if a marker decreases in frequency with age. This information considered and controlled when designing the study. For, example, an age independent marker could be substituted in its place.

25

30

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

EXAMPLE 1

This example describes the use of a database containing information provided by a population of healthy blood donors who were not selected for any particular disease to determine the distribution of allelic frequencies of known genetic markers with age and by sex in a Caucasian subpopulation of the database. The results described in this example demonstrate that a disease-related genetic marker or polymorphism can be identified by sorting a healthy database by a parameter or parameters, such as age, sex and ethnicity.

Generating a database

Blood was obtained by venous puncture from human subjects who met blood bank criteria for donating blood. The blood samples were preserved with EDTA at pH 8.0 and labeled. Each donor provided information such as age, sex, ethnicity, medical history and family medical history. Each sample was labeled with a barcode representing identifying information. A database was generated by entering, for each donor, the subject identifier and information corresponding to that subject into the memory of a computer storage medium using commercially available software, e.g., Microsoft Access.

Model genetic markers

The frequencies of polymorphisms known to be associated at some level with disease were determined in a subpopulation of the subjects represented in the database. These known polymporphisms occur in the p21, p53 and Lipoprotein Lipase genes. Specifically, the N291S polymorphism (N291S) of the Lipoprotein Lipase gene, which results in a substitution of a serine for an asparagine at amino acid codon 291, leads to reduced levels of high density lipoprotein cholesterol (HDL-C) that is associated with an increased risk of males for arteriosclerosis and in particular myocardial infarction (see, Reymer et al. (1995) Nature Genetics 10:28-34).

The p53 gene encodes a cell cycle control protein that assesses DNA damage and acts as a transcription factor regulating genes that control cell of growth, DNA repair and apoptosis (programmed cell death). Mutations in the

15

20

25

30

p53 gene have been found in a wide variety of different cancers, including different types of leukemia, with varying frequency. The loss of normal p53 function results in genomic instability an uncontrolled cell growth. A polymorphism that has been identified in the p53 gene, i.e., the R72P allele, results in the substitution of a proline for an arginine at amino acid codon 72 of the gene.

The p21 gene encodes a cyclin-dependent kinase inhibitor associated with G1 phase arrest of normal cells. Expression of the p21 gene triggers apoptosis. Polymorphisms of the p21 gene have been associated with Wilms' tumor, a pediatric kidney cancer. One polymorphism of the p21 gene, the S31R polymorphism, results in a substitution of an arginine for a serine at amino acid codon 31.

Database analysis

Sorting of subjects according to specific parameters

The genetic polymorphisms were profiled within segments of the Caucasian subpopulation of the sample bank. For p53 profiling, the genomic DNA isolated from blood from a total of 1277 Caucasian subjects age 18-59 years and 457 Caucasian subjects age 60-79 years was analyzed. For p21 profiling, the genomic DNA isolated from blood from a total of 910 Caucasian subjects age 18-49 years and 824 Caucasian subjects age 50-79 years was analyzed. For lipoprotein lipase gene profiling, the genomic DNA from a total of 1464 Caucasian females and 1470 Caucasian males under 60 years of age and a total of 478 Caucasian females and 560 Caucasian males over 60 years of age was analyzed.

Isolation and analysis of genomic DNA

Genomic DNA was isolated from blood samples obtained from the individuals. Ten milliliters of whole blood from each individual was centrifuged at 2000 x g. One milliliter of the buffy coat was added to 9 ml of 155 mM NH₄Cl, 10 mM KHCO₃, and 0.1 mM Na₂EDTA, incubated 10 min at room temperature and centrifuged for 10 min at 2000 x g. The supernatant was removed, and the white cell pellet was washed in 155 mM NH₄Cl, 10 mM KHCO₃ and 0.1 mM Na₂EDTA and resuspended in 4.5 ml of 50 mM Tris, 5 mM

EDTA and 1% SDS. Proteins were precipitated from the cell lysate by 6 mM ammonium acetate, pH 7.3, and then separated from the nucleic acids by centrifugation at 3000 x g. The nucleic acid was recovered from the supernatant by the addition of an equal volume of 100% isopropanol and centrifugation at 2000 x g. The dried nucleic acid pellet was hydrated in 10 mM Tris, pH 7.6, and 1 mM Na₂EDTA and stored at 4° C.

Assays of the genomic DNA to determine the presence or absence of the known genetic markers were developed using the BiomassPROBE™ detection method (primer oligo base extension) reaction. This method uses a single detection primer followed by an oligonucleotide extension step to give products, which can be readily resolved by mass spectrometry, and, in particular, MALDITOF mass spectrometry. The products differ in length depending on the presence or absence of a polymorphism. In this method, a detection primer anneals adjacent to the site of a variable nucleotide or sequence of nucleotides and the primer is extended using a DNA polymerase in the presence of one or more dideoxyNTPs and, optionally, one or more deoxyNTPs. The resulting products are resolved by MALDI-TOF mass spectrometry. The mass of the products as measured by MALDI-TOF mass spectrometry makes possible the determination of the nucleotide(s) present at the variable site.

First, each of the Caucasian genomic DNA samples was subjected to nucleic acid amplification using primers corresponding to sites 5' and 3' of the polymorphic sites of the p21 (S31R allele), p53 (R72P allele) and Lipoprotein Lipase (N291S allele) genes. One primer in each primer pair was biotinylated to permit immobilization of the amplification product to a solid support. Specifically, the polymerase chain reaction primers used for amplification of the relevant segments of the p21, p53 and lipoprotein lipase genes are shown below: US4p21c31-2F (SEQ ID NO: 9) and US5p21-2R (SEQ ID NO: 10) for p21 gene amplification; US4-p53-ex4-F (also shown as p53-ex4US4 (SEQ ID NO: 2)) and US5-p53/2-4R (also shown as US5P53/4R (SEQ ID NO: 3)) for p53 gene amplification; and US4-LPL-F2 (SEQ ID NO: 16) and US5-LPL-R2 (SEQ ID NO: 17) for lipoprotein lipase gene amplification.

PCT/US00/28413

20

25

30

-44-

Amplification of the respective DNA sequences was conducted according to standard protocols. For example, primers may be used in a concentration of 8 pmol. The reaction mixture (e.g., total volume 50 μ l) may contain Taq-polymerase including 10x buffer and dTNPs. Cycling conditions for polymerase chain reaction amplification may typically be initially 5 min. at 95°C, followed by 1 min. at 94°C, 45 sec at 53°C, and 30 sec at 72°C for 40 cycles with a final extension time of 5 min at 72°C. Amplification products may be purified by using Qiagen's PCR purification kit (No. 28106) according to manufacturer's instructions. The elution of the purified products from the column can be done in 50 μ l TE-buffer (10mM Tris, 1 mM EDTA, pH 7.5).

The purified amplification products were immobilized via a biotin-avidin linkage to streptavidin-coated beads and the double-stranded DNA was denatured. A detection primer was then annealed to the immobilized DNA using conditions such as, for example, the following: $50 \,\mu$ l annealing buffer (20 mM Tris, 10 mM KCl, 10 mM (NH₄)₂SO₄, 2 mM MgSO₂, 1% Triton X-100, pH 8) at 50° C for 10 min, followed by washing of the beads three times with 200 μ l washing buffer (40 mM Tris, 1 mM EDTA, 50 mM NaCl, 0.1% Tween 20, pH 8.8) and once in 200 μ l TE buffer.

The PROBE extension reaction was performed, for example, by using some components of the DNA sequencing kit from USB (No. 70770) and dNTPs or ddNTPs from Pharmacia. An exemplary protocol could include a total reaction volume of 45 μ l, containing of 21 μ l water, 6 μ l Sequenase-buffer, 3 μ l 10 mM DTT solution, 4.5 μ l, 0.5 mM of three dNTPs, 4.5 μ l, 2 mM the missing one ddNTP, 5.5 μ l glycerol enzyme dilution buffer, 0.25 μ l Sequenase 2.0, and 0.25 pyrophosphatase. The reaction can then by pipetted on ice and incubated for 15 min at room temperature and for 5 min at 37°C. The beads may be washed three times with 200 μ l washing buffer and once with 60 μ l of a 70 mM NH₄-Citrate solution.

The DNA was denatured to release the extended primers from the immobilized template. Each of the resulting extension products was separately analyzed by MALDI-TOF mass spectrometry using 3-hydroxypicolinic acid (3-HPA) as matrix and a UV laser.

WO 01/27857 PCT/US00/28413

-45-

Specifically, the primers used in the PROBE reactions are as shown below: P21/31-3 (SEQ ID NO: 12) for PROBE analysis of the p21 polymorphic site; P53/72 (SEQ ID NO: 4) for PROBE analysis of the p53 polymorphic site; and LPL-2 for PROBE analysis of the lipoprotein lipase gene polymorphic site. In the PROBE analysis of the p21 polymorphic site, the extension reaction was performed using dideoxy-C. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 31 encodes a serine) and from the reaction conducted on a polymorphic S31R allele template (wherein codon 31 encodes an arginine) are shown below and designated as P21/31-3 Ser (wt) (SEQ ID NO: 13) and P21/31-3 Arg (SEQ ID NO: 14), respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 4900.2 Da for the wild-type product and 5213.4 Da for the polymorphic product).

In the PROBE analysis of the p53 polymorphic site, the extension reaction was performed using dideoxy-C. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 72 encodes an arginine) and from the reaction conducted on a polymorphic R72P allele template (wherein codon 72 encodes a proline) are shown below and designated as Cod72 G Arg (wt) and Cod72 C Pro, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 5734.8 Da for the wild-type product and 5405.6 Da for the polymorphic product).

In the PROBE analysis of the lipoprotein lipase gene polymorphic site, the extension reaction was performed using a mixture of ddA and ddT. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 291 encodes an asparagine) and from the reaction conducted on a polymorphic N291S allele template (wherein codon 291 encodes a serine) are shown below and designated as 291Asn and 291Ser, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 6438.2 Da for the wild-type product and 6758.4 Da for the polymorphic product).

P53-1 (R72P)

25

PCR Product length: 407 bp (SEQ ID NO: 1)

Primers (SEQ ID NOs: 2-4)

p53-ex4FUS4 ccc agt cac gac gtt gta aaa cgc tga gga cct ggt cct ctg ac

15 US5P53/4R

age qga taa caa ttt cac aca qqt tga agt ctc atg gaa gcc

P53/72

20.

35

gcc aga ggc tgc tcc cc

Masses

Allele	Product Termination: ddC	SEQ #	Length	Mass
P53/72	gccagaggctgctcccc	5	17	5132.4
Cod72 G Arg (wt)	gccagaggctgctccccgc	6	19	5734.8
Cod72 C Pro	gccagaggctgctccccc	7	18	5405.6

Biotinylated US5 primer is used in the PCR amplification.

LPL-1 (N291S)

Amino acid exchange asparagine to serine at codon 291 of the lipoprotein lipase gene.

PCR Product length: 251 bp (SEQ ID NO: 15)

US4-LPL-F2 (SEQ ID NO: 16)

gcgctccatt catctctca tcgactctct gttgaatgaa gaaaatccaa gtaaggccta
caggtgcagt tccaaggaag cctttgagaa agggctctgc ttgagttgta gaaagaaccg
LPL-2 291N
ctgcaa<u>caat ctqqqctatq aqatca</u>ataa agtcagagcc aaaagaagca gcaaaatgta
g 291S

cctgaagact cgttctcaga tgccc US4-LPL-R2

Primers (SEQ ID NOs: 16-18):

US4-LPL-F2 ccc agt cac gac gtt gta aaa cgg cgc tcc att cat ctc ttc
US5-LPL-R2 agc gga taa caa ttt cac aca ggg ggc atc tga gaa cga gtc
LPL-2 caa tct ggg cta tga gat ca

Masses

Allele	Product Termination: ddA, ddT	SEQ #	Length	Mass
LPL-2	caatctgggctatgagatca	19	20	6141
291 Asn	caatctgggctatgagatcaa	20	21	6438.2
291 Ser	caatctgggctatgagatcagt	21	22	6758.4

Biotinylated US5 primer is used in the PCR amplification.

P21-1 (S31R)

Amino acid exchange serine to arginine at codon 31 of the tumor suppressor gene p21. Product length: 207 bp (SEQ ID NO: 8) US4p21c31-2F

gtcc gtcagaaccc atgcggcagc p21/31-3 31S

aaggeetgee geegeetett eggeeeagtg ga<u>caqeqaqe aqetqaq</u>eeg egaetgtgat a 31R gegetaatgg egggetgeat ecaggaggee egtgagegat ggaacttega etttgteace gagacaccac tggaggg US5p21-2R

Primers (SEQ ID NOs: 9-11)

20 US4p21c31-2F ccc agt cac gac gtt gta aaa cgg tcc gtc aga acc cat gcg g

US5p21-2R agc gga taa caa ttt cac aca ggc tcc agt ggt gtc tcg gtg ac

P21/31-3 cag cga gca gct gag

Masses

15

25

30

35

Allele	Product Termination: ddC	SEQ #	Length	Mass
p21/31-3	cagcgagcagctgag	12	15	4627
P21/31-3 Ser (wt)	cagcgagcagctgagc	13	16	4900.2
P21/31-3 Arg	cagcgagcagctgagac	. 14	17	5213.4

Biotinylated US5 primer is used in the PCR amplification.

Each of the Caucasian subject DNA samples was individually analyzed by MALDI-TOF mass spectrometry to determine the identity of the nucleotide at the polymorphic sites. The genotypic results of each assay can be entered into the database. The results were then sorted according to age and/or sex to determine the distribution of allelic frequencies by age and/or sex. As depicted in the Figures showing

25

histograms of the results, in each case, there was a differential distribution of the allelic frequencies of the genetic markers for the p21, p53 and lipoprotein lipase gene polymorphisms.

Figure 8 shows the results of the p21 genetic marker assays reveals a statistically significant decrease (from 13.3% to 9.2%) in the frequency of the heterozygous genotype (S31R) in Caucasians with age (18-49 years of age compared to 50-79 years of age). The frequencies of the homozygous (S31 and R31) genotypes for the two age groups are also shown, as are the overall frequencies of the S31 and R31 alleles in the two age groups (designated as 'S31 and 'R31, respectively in the Figure).

Figures 7A-C shows the results of the p53 genetic marker assays and reveals a statistically significant decrease (from 6.7% to 3.7%) in the frequency of the homozygous polymorphic genotype (P72) in Caucasians with age (18-59 years of age compared to 60-79 years of age). The frequencies of the homozygous "wild-type" genotype (R72) and the heterozygous genotype (R72P) for the two age groups are also shown, as are the overall frequencies of the R72 and P72 alleles in the two age groups (designated as 'R72 and 'P72, respectively in the Figure). These results are consistent with the observation that allele is not benign, as p53 regulates expression of a second protein, p21, which inhibits cyclin-dependent kinases (CDKs) needed to drive cells through the cell-cycle (a mutation in either gene can disrupt the cell cycle leading to increased cell division).

Figure 2C shows the results of the lipoprotein lipase gene genetic marker assays reveals a statistically significant decrease (from 1.97% to 0.54%) in the frequency of the polymorphic allele (S291) in Caucasian males with age (see also Reymer et al. (1995) Nature Genetics 10:28-34).

PCT/US00/28413

-49-

The frequencies of this allele in Caucasian females of different age groups are also shown.

EXAMPLE 2

This example describes the use of MALDI-TOF mass spectrometry to analyze DNA samples of a number of subjects as individual samples and as pooled samples of multiple subjects to assess the presence or absence of a polymorphic allele (the 353Q allele) of the Factor VII gene and determine the frequency of the allele in the group of subjects. The results of this study show that essentially the same allelic frequency can be obtained by analyzing pooled DNA samples as by analyzing each sample separately and thereby demonstrate the quantitative nature of MALDI-TOF mass spectrometry in the analysis of nucleic acids.

Factor VII

Factor VII is a serine protease involved in the extrinsic blood coagulation cascade. This factor is activated by thrombin and works with tissue factor (Factor III) in the processing of Factor X to Factor Xa. There is evidence that supports an association between polymorphisms in the Factor VII gene and increased Factor VII activity which can result in an elevated risk of ischemic cardiovascular disease, including myocardial infarction. The polymorphism investigated in this study is R353Q (i.e., a substitution of a glutamic acid residue for an arginine residue at codon 353 of the Factor VII gene) (see Table 5).

Analysis of DNA samples for the presence or absence of the 353Q allele of the Factor VII gene

25

Genomic DNA was isolated from separate blood samples obtained from a large number of subjects divided into multiple groups of 92 subjects per group. Each sample of genomic DNA was analyzed using the BiomassPROBE™ assay as described in Example 1 to determine the presence or absence of the 353Q polymorphism of the Factor VII gene:

First, DNA from each sample was amplified in a polymerase chain reaction using primers F7-353FUS4 (SEQ ID NO: 24) and F7-353RUS5 (SEQ ID NO: 26) as shown below and using standard conditions, for example, as described in Example 1. One of the primers was biotinylated to permit immobilization of the amplification product to a solid support. The purified amplification products were immobilized via a biotin-avidin linkage to streptavidin-coated beads and the double-stranded DNA was denatured. A detection primer was then annealed to the immobilized DNA using conditions such as, for example, described in Example 1. The detection primer is shown as F7-353-P (SEQ ID NO: 27) below. The PROBE extension reaction was carried out using conditions, for example, such as those described in Example 1. The reaction was performed using ddG.

The DNA was denatured to release the extended primers from the immobilized template. Each of the resulting extension products was separately analyzed by MALDI-TOF mass spectrometry. A matrix such as 3-hydroxypicolinic acid (3-HPA) and a UV laser could be used in the MALDI-TOF mass spectrometric analysis. The products resulting from the reaction conducted on a "wild-type" allele template (wherein codon 353 encodes an arginine) and from the reaction conducted on a polymorphic 353Q allele template (wherein codon 353 encodes a glutamic acid) are shown below and designated as 353 CGG and 353 CAG, respectively. The masses for each product as can be measured by MALDI-TOF mass spectrometry are also provided (i.e., 5646.8 Da for the wild-type product and 5960 Da for the polymorphic product).

The MALDI-TOF mass spectrometric analyses of the PROBE reactions of each DNA sample were first conducted separately on each sample (250 nanograms total concentration of DNA per analysis). The allelic frequency of the 353Q polymorphism in the group of 92 subjects

was calculated based on the number of individual subjects in which it was detected.

Next, the samples from 92 subjects were pooled (250 nanograms total concentration of DNA in which the concentration of any individual DNA is 2.7 nanograms) and the pool of DNA was subjected to MALDITOF mass spectrometric analysis. The area under the signal corresponding to the mass of the 353Q polymorphism PROBE extension product in the resulting spectrum was integrated in order to quantitate the amount of DNA present. The ratio of this amount to total DNA was used to determine the allelic frequency of the 353Q polymorphism in the group of subjects. This type of individual sample vs. pooled sample analysis was repeated for numerous different groups of 92 different samples.

The frequencies calculated based on individual MALDI-TOF mass spectrometric analysis of the 92 separate samples of each group of 92 are compared to those calculated based on MALDI-TOF mass spectrometric analysis of pools of DNA from 92 samples in Figure 9. These comparisons are shown as "pairs" of bar graphs in the Figure, each pair being labeled as a separate "pool" number, e.g., P1, P16, P2, etc. Thus, for example, for P1, the allelic frequency of the polymorphism calculated by separate analysis of each of the 92 samples was 11.41% and the frequency calculated by analysis of a pool of all of the 92 DNA samples was 12.09%.

The similarity in frequencies calculated by analyzing separate DNA samples individually and by pooling the DNA samples demonstrates that it is possible, through the quantitative nature of MALDI-TOF mass spectrometry, to analyze pooled samples and obtain accurate frequency determinations. The ability to analyze pooled DNA samples significantly reduces the time and costs involved in the use of the non-selected, healthy databases as described herein. It has also been shown that it is

64°C

possible to decrease the DNA concentration of the individual samples in a pooled mixture from 2.7 nanograms to 0.27 nanograms without any change in the quality of the spectrum or the ability to quantitate the amount of sample detected.

5 Factor VII R353Q PROBE Assay

PROBE Assay for cod353 CGG>CAG (Arg>Gln), Exon 9 G>A.

PCR fragment: 134 bp (incl. US tags; SEQ ID Nos. 22 and 23)

Frequency of A allele: Europeans about 0.1, Japanese/Chinese about 0.03-0.05 (Thromb. Haemost. 1995, 73:617-22; Diabetologia 1998,

10 41:760-6):

F7-353FUS4>

1201 GTGCCGGCTA CTCGGATGGC AGCAAGGACT CCTGCAAGGG GGACAGTGGA
GGCCCACATG
F7-353-P> A <F7-353RUS5

F7-353-P> A <F7-353RUS5

15 1261 CCACCACTA CCGGGGCACG TGGTACCTGA CGGGCATCGT CAGCTGGGGC

CAGGGCTGCG

Primers (SEQ ID NOs: 24-26)

F7-353FUS4 CCC AGT CAC GAC GTT GTA AAA CGA TGG CAG CAA GGA CTC CTG

F7-353-P CAC ATG CCA CCC ACT ACC

20 F7-353RUS5 AGC GGA TAA CAA TTT CAC ACA GGT GAC GAT GCC CGT CAG GTA C 6

Masses

25

Allele	Product Termination: ddG	SEQ #	Length	Mass
F7-353-P	atgccaccactacc	27	18	5333.6
353 CGG	cacatgccacccactaccg	28	19	5646.8
353 CAG	cacatgccacccactaccag	29	20	5960
US5-bio bio-	agcggataacaatttcacacagg	30	23	7648.6

Conclusion

The above examples demonstrate an effect of altered frequency of disease causing genetic factors within the general population.

Interpretation of those results allows prediction of the medical relevance of polymorphic genetic alterations. In addition, conclusions can be drawn with regard to their penetrance, diagnostic specificity, positive predictive value, onset of disease, most appropriate onset of preventive strategies,

20

and the general applicability of genetic alterations identified in isolated populations to panmixed populations. Therefore, an age- and sex-stratified population-based sample bank that is ethnically homogenous is a suitable tool for rapid identification and validation of genetic factors regarding their potential medical utility.

EXAMPLE 3

MORBIDITY AND MORTALITY MARKERS

Sample Band and Initial Screening

Healthy samples were obtained through the blood bank of San

10 Bernardino, CA. Donors signed prior to the blood collection a consent form and agreed that their blood will be used in genetic studies with regard to human aging. All samples were anomymized. Tracking back of samples is not possible.

Isolation of DNA from blood samples of a healthy donor population

Blood is obtained from a donor by venous puncture and preserved with 1mM EDTA pH 8.0. Ten milliliters of whole blood from each donor was centrifuged at 2000x g. One milliliter of the buffy coat was added to 9 milliters of 155mM NH₄Cl, 10mM KHCO₃, and 0.1mM Na₂EDTA, incubated 10 minutes at room temperature and centrifuged for 10 minutes at 2000x g. The supernatant was removed, and the white cell pellet was washed in 155mM NH₄Cl, 10mM KHCO₃, and 0.1mM Na₂EDTA and resuspended in 4.5 milliliters of 50mM Tris, 5mM EDTA, and 1% SDS. Proteins were precipitated from the cell lysate by 6M Ammonium Acetate, pH 7.3, and separated from the nucleic acid by centrifugation 3000x g. The nucleic acid was recovered from the supernatant by the addition of an equal volume of 100% isopropanol and centrifugation at 2000x g. The dried nucleic acid pellet was hydrated in lOmM Tris pH 7.6 and 1mM Na2EDTA and stored at 4C.

In this study, samples were pooled as shown in Table 1. Both parents of the blood donors were of Caucasian origin.

Table 1

Pool ID	Sex	Age-range	# individuals
SP1	Female	18-39 years	276
SP2	Males	18-39 years	276
SP3	Females	60-69 years	184
SP4	Males	60-79 years	368

More than 400 SNPs were tested using all four pools. After one test run 34 assays were selected to be re-assayed at least once. Finally, 10 assays showed repeatedly differences in allele frequencies of several percent and, therefore, fulfilled the criteria to be tested using the individual samples. Average allele frequency and standard deviation is tabulated in Table 2.

Table 2

SP1	SP1-STD	SP2	SP2-STD	SP3	SP3-STD	SP4	SP4-STD
0.457	0.028	0.433	0.042	0.384	0.034	0.380	0.015
0.276	0.007	0.403	0.006	0.428	0.052	0.400	0.097
0.676	0.013	0.627	0.018	0.755	0.009	0.686	0.034
0.581	0.034	0.617	0.045	0.561	n.a.	0.539	0.032
0.504	0.034	0.422	0.020	0.477	0.030	0.556	0.005
0.537	0.017	0.503	n.a.	0.623	0.023	0.535	0.009
0.560	0.006	0.527	0.059	0.546	0.032	0.436	0.016
0.754	0.008	0.763	0.047	0.736	0.052	0.689	0.025
0.401	0.022	0.363	0.001	0.294	0.059	0.345	0.01.3
	0.457 0.276 0.676 0.581 0.504 0.537 0.560 0.754	0.457 0.028 0.276 0.007 0.676 0.013 0.581 0.034 0.504 0.034 0.537 0.017 0.560 0.006 0.754 0.008	0.457 0.028 0.433 0.276 0.007 0.403 0.676 0.013 0.627 0.581 0.034 0.617 0.504 0.034 0.422 0.537 0.017 0.503 0.560 0.006 0.527 0.754 0.008 0.763	0.457 0.028 0.433 0.042 0.276 0.007 0.403 0.006 0.676 0.013 0.627 0.018 0.581 0.034 0.617 0.045 0.504 0.034 0.422 0.020 0.537 0.017 0.503 n.a. 0.560 0.006 0.527 0.059 0.754 0.008 0.763 0.047	0.457 0.028 0.433 0.042 0.384 0.276 0.007 0.403 0.006 0.428 0.676 0.013 0.627 0.018 0.755 0.581 0.034 0.617 0.045 0.561 0.504 0.034 0.422 0.020 0.477 0.537 0.017 0.503 n.a. 0.623 0.560 0.006 0.527 0.059 0.546 0.754 0.008 0.763 0.047 0.736	0.457 0.028 0.433 0.042 0.384 0.034 0.276 0.007 0.403 0.006 0.428 0.052 0.676 0.013 0.627 0.018 0.755 0.009 0.581 0.034 0.617 0.045 0.561 n.a. 0.504 0.034 0.422 0.020 0.477 0.030 0.537 0.017 0.503 n.a. 0.623 0.023 0.560 0.006 0.527 0.059 0.546 0.032 0.754 0.008 0.763 0.047 0.736 0.052	SPI. SPI.SIB SPZ SPZ SPZ-SIB D.380 D.380

20

25

	63306	0.697	0.012	0.674	0.013	0.712	0.017	0.719	0.005
--	-------	-------	-------	-------	-------	-------	-------	-------	-------

So far, 7 out of the 10 potential morbidity markers were fully analyzed. Additional information about genes in which these SNPs are located was gathered through publicly databases like Genbank.

AKAPS

20

Candidate morbidity and mortality markers include housekeeping genes, such as genes involved in signal transduction. Among such genes are the A-kinase anchoring proteins (AKAPs) genes, which participate in signal transduction pathways involving protein phosphorylation. Protein phosphorylation is an important mechanism for enzyme regulation and the transduction of extracellular signals across the cell membrane in eukaryotic cells. A wide variety of cellular substrates, including enzymes, membrane receptors, ion channels and transcription factors, can be phosphorylated in response to extracellular signals that interact with cells. A-key enzyme in the phosphorylation of cellular proteins in response to hormones and neurotransmitters is cyclic AMP (cAMP)-dependent protein kinase (PKA). Upon activation by cAMP, PKA thus mediates a variety of cellular responses to such extracellular signals. An array of PKA isozymes are expressed in mammalian cells. The PKAs usually exist as inactive tetramers containing a regulatory (R) subunit dimer and two catalytic (C) subunits. Genes encoding three C subunits (Ca, C β and Cy) and four R subunits (RI α , RI β , RII α and RII β) have been identified [see Takio et al. (1982) Proc. Natl. Acad. Sci. U.S. A. 79:2544-2548; Lee et al. (1983) Proc. Natl. Acad. Sci. U.S. A. 80:3608-3612; Jahnsen et al. (1996) J. Biol. Chem. 261:12352-12361; Clegg et al. (1988) Proc. Natl. Acad. Sci. U.S. A. 85:3703-3707; and Scott (1991) Pharmacol. Ther. 50:123-145]. The type I (RI) α and type II (RII) α subunits are distributed ubiquitously, whereas $RI\beta$ and $RII\beta$ are present mainly in brain [see. e.g., Miki and Eddy

(1999) J. Biol. Chem. 274:29057-29062]. The type I PKA holoenzyme (RI α and RI $oldsymbol{eta}$) is predominantly cytoplasmic, whereas the majority of type II PKA (RIIa and RII $oldsymbol{eta}$) associates with cellular structures and organelles [Scott (1991) Pharmacol. Ther. 50:123-145]. Many hormones and other signals act through receptors to generate cAMP which binds to the R subunits of PKA and releases and activates the C subunits to phosphorylate proteins. Because protein kinases and their substrates are widely distributed throughout cells, there are mechanisms in place in cells to localize protein kinase-mediated responses to different signals. One such mechanism involves subcellular targeting of PKAs through association with anchoring proteins, referred to as A-kinase anchoring proteins (AKAPs), that place PKAs in close proximity to specific organelles or cytoskeletal components and particular substrates thereby providing for more specific PKA interactions and localized responses [see, e.g., Scott et al. (1990) J. Biol. Chem. 265:21561-21566; Bregman et al. (1991) J. Biol. Chem. 266:7207-7213; and Miki and Eddy (1999) J. Biol. Chem. 274:29057-29062]. Anchoring not only places the kinase close to preferred substrates, but also positions the PKA holoenzyme at sites where it can optimally respond to fluctuations in the second messenger cAMP [Mochly-Rosen (1995) Science 268:247-251; Faux and Scott (1996) Trends Biochem. Sci. 21:312-315; Hubbard and Cohen (1993) Trends Biochem. Sci. 18:172-177].

Up to 75% of type II PKA is localized to various intracellular sites through association of the regulatory subunit (RII) with AKAPs [see, e.g., Hausken et al. (1996) J. Biol. Chem. 271:29016-29022]. RII subunits of PKA bind to AKAPs with nanomolar affinity [Carr et al. (1992) J. Biol. Chem. 267:13376-13382], and many AKAP-RII complexes have been isolated from cell extracts. RI subunits of PKA bind to AKAPs with only micromolar affinity [Burton et al. (1997) Proc. Natl. Acad. Sci. U.S.A.

94:11067-11072]. Evidence of binding of a PKA RI subunit to an AKAP has been reported [Miki and Eddy (1998) *J. Biol. Chem 273*:34384-34390] in which RIa-specific and RIa/RIIa dual specificity PKA anchoring domains were identified on FSC1/AKAP82. Additional dual specific AKAPs, referred to as D-AKAP1 and D-AKAP2, which interact with the type I and type II regulatory subunits of PKA have also been reported [Huang et al. (1997) *J. Biol. Chem. 272*:8057-8064; Huang et al. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*:11184-11189].

More than 20 AKAPs have been reported in different tissues and species. Complementary DNAs (cDNAs) encoding AKAPs have been 10 isolated from diverse species, ranging from Caenorhabditis elegans and Drosophilia to human [see, e.g., Colledge and Scott (1999) Trends Cell Biol. 9:216-221]. Regions within AKAPs that mediate association with RII subunits of PKA have been identified. These regions of approximately 10-18 amino acid residues vary substantially in primary sequence, but 15 secondary structure predictions indicate that they are likely to form an amphipathic helix with hydrophobic residues aligned along one face of the helix and charged residues along the other [Carr et al. (1991) J. Biol. Chem. 266:14188-14192; Carr et al. (1992) J. Biol. Chem. 267:13376-13382]. Hydrophobic amino acids with a long aliphatic side chain, e.g., 20 valine, leucine or isoleucine, may participate in binding to RII subunits [Glantz et al. (1993) J. Biol. Chem. 268:12796-12804].

Many AKAPs also have the ability to bind to multiple proteins, including other signaling enzymes. For example, AKAP79 binds to PKA, protein kinase C (PKC) and the protein phosphatase calcineurin (PP2B) [Coghlan et al. (1995) Science 267:108-112 and Klauck et al. (1996) Science 271:1589-1592]. Therefore, the targeting of AKAP79 to neuronal postsynaptic membranes brings together enzymes with opposite catalytic activities in a single complex.

AKAP10

20

AKAPs thus serve as potential regulatory mechanisms that increase the selectivity and intensity of a cAMP-mediated response. There is a need, therefore, to identify and elucidate the structural and functional properties of AKAPs in order to gain a complete understanding of the important role these proteins play in the basic functioning of cells.

The sequence of a human AKAP10 cDNA (also referred to as D-AKAP2) is available in the GenBank database, at accession numbers AF037439 (SEQ ID NO: 31) and NM 007202. The AKAP10 gene is located on chromosome 17.

The sequence of a mouse D-AKAP2 cDNA is also available in the GenBank database (see accession number AF021833). The mouse D-AKAP2 protein contains an RGS domain near the amino terminus that is characteristic of proteins that interact with Ga subunits and possess GTPase activating protein-like activity [Huang et al. (1997) Proc. Natl. Acad. Sci. U.S.A. 94:11184-11189]. The human AKAP10 protein also has sequences homologous to RGS domains. The carboxy-terminal 40 residues of the mouse D-AKAP2 protein are responsible for the interaction with the regulatory subunits of PKA. This sequence is fairly well conserved between the mouse D-AKAP2 and human AKAP10 proteins. Polymorphisms of the human AKAP10 gene and polymorphic AKAP10 proteins

Polymorphisms of AKAP genes that alter gene expression, regulation, protein structure and/or protein function are more likely to have a significant effect on the regulation of enzyme (particularly PKA) activity, cellular transduction of signals and responses thereto and on the basic functioning of cells than polymorphisms that do not alter gene and/or protein function. Included in the polymorphic AKAPs provided herein are human AKAP10 proteins containing differing amino acid residues at position number 646.

30

Amino acid 646 of the human AKAP10 protein is located in the carboxy-terminal region of the protein within a segment that participates in the binding of R-subunits of PKAs. This segment includes the carboxy-terminal 40 amino acids.

The amino acid residue reported for position 646 of the human AKAP10 protein is an isoleucine. Polymorphic human AKAP10 proteins provided herein have the amino acid sequence but contain residues other than isoleucine at amino acid position 646 of the protein. In particular embodiments of the polymorphic human AKAP10 proteins provided herein, the amino acid at position 646 is a valine, leucine or phenylalanine residue.

An A to G transition at nucleotide 2073 of the human AKAP10 coding sequence

As described herein, an allele of the human AKAP10 gene that contains a specific polymorphism at position 2073 of the coding sequence and thereby encodes a valine at position 646 has been detected in varying frequencies in DNA samples from younger and older segments of the human population. In this allele, the A at position 2073 of the AKAP10 gene coding sequence is changed from an A to a G, giving rise to an altered sequence in which the codon for amino acid 646 changes from ATT, coding for isoleucine, to GTT, coding for valine.

Morbidity marker 1: human protein kinase A anchoring protein (AKAP10-1)

PCR Amplification and BiomassPROBE assay detection of AKAP10-1 in a healthy donor population

PCR Amplification of donor population for AKAP 10

PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out in single 50µl PCR reaction with 100ng-1ug of pooled human genomic DNAs in a 50µl PCR reaction. Individual DNA concentrations within the

pooled samples were present in equal concentration with the final concentration ranging from 1-25ng. Each reaction containing IX PCR buffer (Qiagen, Valencia, CA), 200uM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl₂, and 25pmol of the 5 forward primer containing the universal primer sequence and the target specific sequence 5'-TCTCAATCATGTGCATTGAGG-3'(SEQ ID NO: 45), 2pmol of the reverse primer 5'-AGCGGATAACAATTTCACACAGGGATCACACAGCCATCAGCAG-3' (SEQ ID NO: 46), and lOpmol of a biotinylated universal primer complementary to the 5' end of the PCR amplicon 5'-AGCGGATAACAATTTCACACAGG-3'(SEQ ID NO: 47). After an initial round of amplification with the target with the specific forward and reverse primer, the 5' biotinylated universal primer then hybridized and acted as a reverse primer thereby introducing a 3' biotin capture moiety into the molecule. The amplification protocol results in a 5'-biotinylated double stranded DNA amplicon and dramatically reduces the cost of high throughput genotyping by eliminating the need to 5' biotin label each forward primer used in a genotyping. Thermal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; 72° C 3min.

Immobilization of DNA

The 50µl PCR reaction was added to 25ul of streptavidin coated magnetic bead (Dynal) prewashed three times and resuspended in 1M NH₄Cl, 0.06M NH₄OH. The PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The unbound strand was release from the double stranded amplicons by

incubation in 100mM NaOH and washing of the beads three times with 10mM Tris pH 8.0.

BiomassPROBE assay analysis of donor population for AKAP10-1 (clone 48319)

Genotyping using the BiomassPROBE assay methods was carried out by resuspending the DNA coated magnetic beads in 26mM Tris-HCI pH 9.5, 6.5 mM MgCl $_{
m 2}$ and 50mM each of dTTP and 50mM each of ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase (Ambersham) and 20pmol of a template specific oligonucleotide PROBE primer 5'-CTGGCGCCCACGTGGTCAA-3' (SEQ ID NO: 48) (Operon). Primer extension occurs with three cycles of oligonucleotide primer hybridization and extension. The extension products were analyzed after denaturation from the template with 50mM NH₄Cl and transfer of 150nL each sample to a silicon chip preloaded with 150nL of H3PA matrix 15 material. The sample material was allowed to crystallize and was analyzed by MALDI-TOF (Bruker, PerSeptive). The SNP that is present in AKAP10-1 is a T to C transversion at nucleotide number 156277 of the sequence of a genomic clone of the AKAP10 gene (GenBank Accession No. AC005730) (SEQ ID NO: 36). SEQ ID NO: 35: represents the nucleotide sequence of human chromosome 17, which contains the genomic nucleotide sequence of the human AKAP10 gene, and SEQ ID NO: represents the nucleotide sequence of human chromosome 17, which contains the genomic nucleotide sequence of the human AKAP10-1 allele. The mass of the primer used in the BioMass probe reaction was 5500.6 daltons. In the presence of the SNP, the primer is extended by the addition of ddC, which has a mass of 5773.8. The wildtype gene results in the addition of dT and ddG to the primer to produce an extension product having a mass of 6101 daltons.

20

25

The frequency of the SNP was measured in a population of age selected healthy individuals. Five hundred fifty-two (552) individuals between the ages of 18-39 years (276 females, 276 males) and 552 individuals between the ages of 60-79 (184 females between the ages of 60-69, 368 males between the age of 60-79) were tested for the presence of the polymorphism localized in the non-translated 3'region of AKAP 10. Differences in the frequency of this polymorphism with increasing age groups were observed among healthy individuals. Statistical analysis showed that the significance level for differences in the allelic frequency for alleles between the "younger" and the "older" populations was p=0.0009 and for genotypes was p=0.003. Differences between age groups are significant. For the total population allele significance is p=0.0009, and genotype significance is p=0.0003.

This marker led to the best significant result with regard to allele and genotype frequencies in the age-stratified population. Figure 19 shows the allele and genotype frequency in both genders as well as in the entire population. For latter the significance for alleles was $p\!=\!0.0009$ and for genotypes was $p\!=\!0.003$. The young and old populations were in Hardy-Weinberg equilibrium. A preferential change of one particular genotype was not seen.

The polymorphism is localized in the non-translated 3'-region of the gene encoding the human protein kinase A anchoring protein (AKAP10). The gene is located on chromosome 17. Its structure includes 15 exons and 14 intervening sequences (introns). The encoded protein is responsible for the sub-cellular localization of the cAMP-dependent protein kinase and, therefore, plays a key role in the G-protein mediated receptor-signaling pathway (Huang et al. PNAS (1007) 94:11184-11189). Since its localization is outside the coding region, this polymorphism is most likely in linkage disequilibrium (LD) with other non-synonymous

polymorphisms that could cause amino acid substitutions and subsequently alter the function of the protein. Sequence comparison of different Genbank database entries concerning this gene revealed further six potential polymorphisms of which two are supposed to change the respective amino acid (see Table 3).

Table 3

10

20

25

30

Exon	Codon	Nucleotides	Amino acid
3	100	GCT>GCC	Ala > Ala
4	177	AGT > GTG	Met>Val
8	424	GGG > GGC	Gly > Gly
10	524	CCG > CTG	Pro > Leu
12	591	GTG > GTC	Val > Val
12	599	CGC > CGA	Arg > Arg

15 Morbitity marker 2: human protein kinase A anchoring protein (AKAP10-5)

Discovery of AKAP10-5 Allele (SEQ ID NO: 33)

Genomic DNA was isolated from blood (as described above) of seventeen (17) individuals with a genotype CC at the AKAP10-1 gene locus and a single heterozygous individual (CT) (as described). A target sequence in the AKAP10-1 gene which encodes the C-terminal PKA binding domain was amplified using the polymerase chain reaction. PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10-1 target sequence was carried out in individual 50µl PCR reaction with 25ng of human genomic DNA templates. Each reaction containing I X PCR buffer (Qiagen, Valencia, CA), 200µM dNTPs, IU Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl₂, 25pmol of the forward primer (Ex13F) containing the universal primer sequence and the target specific sequence 5′-TCC CAA AGT GCT GGA ATT AC-3′ (SEQ ID NO: 53), and 2pmol of the reverse

20

primer (Ex14R) 5'-GTC CAA TAT ATG CAA ACA GTT G-3' (SEQ ID NO: 54). Thermal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler (MJ Research, Waltham, MA) (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles; 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; 72° C 3min. After amplification the amplicons were purified using a chromatography (Mo Bio Laboratories (Solana Beach, CA)).

The sequence of the 18 amplicons, representing the target region, was determined using a standard Sanger cycle sequencing method with 25nmol of the PCR amplicon, 3.2uM DNA sequencing primer 5'-CCC ACA GCA GTT AAT CCT TC-3'(SEQ ID NO: 55), and chain terminating dRhodamine labeled 2', 3' dideoxynucleotides (PE Biosystems, Foster City, CA) using the following cycling parameters: 96° C for 15 seconds; 25 cycles: 55° C for 15 seconds, 60° C for 4 minutes. The sequencing products precipitated by 0.3M NaOAc and ethanol. The precipitate was centrifuged and dried. The pellets were resuspended in deionized formamide and separated on a 5% polyacrylimide gel. The sequence was determined using the "Sequencher" software (Gene Codes, Ann Arbor, MI).

The sequence of all 17 of the amplicons, which are homozygous for the AKAP10-1 SNP of the amplicons, revealed a polymorphism at nucleotide position 152171 (numbering for GenBank Accession No. AC005730 for AKAP10 genomic clone (SEQ ID NO: 35)) with A replaced by G. This SNP can also be designated as located at nucleotide 2073 of a cDNA clone of the wildtype AKAP10 (GenBank Accession No. 25 AF037439) (SEQ ID NO: 31). The amino acid sequence of the human AKAP10 protein is provided as SEQ ID NO: 32. This single nucleotide polymorphism was designated as AKAP10-5 (SEQ ID NO: 33) and resulted in a substitution of a valine for an isoleucine residue at amino-

20

acid position 646 of the amino acid sequence of human AKAP10 (SEQ ID NO: 32).

PCR Amplification and BiomassPROBE assay detection of AKAP10-5 in a healthy donor population

The healthy population stratified by age is a very efficient and a universal screening tool for morbidity associated genes by allowing for the detection of changes of allelic frequencies in the young compared to the old population. Individual samples of this healthy population base can be pooled to further increase the throughput.

Healthy samples were obtained through the blood bank of San Bernardino, CA. Both parents of the blood donors were of Caucasian origin. Practically a healthy subject, when human, is defined as human donor who passes blood bank criteria to donate blood for eventual use in the general population. These criteria are as follows: free of detectable viral, bacterial, mycoplasma, and parasitic infections; not anemic; and then further selected based upon a questionnaire regarding history (see Figure 3). Thus, a healthy population represents an unbiased population of sufficient health to donate blood according to blood bank criteria, and not further selected for any disease state. Typically such individuals are not taking any medications.

PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out in a single 50µl PCR reaction with 100ng- 1µg of pooled human genomic DNAs in a 50µl PCR reaction. Individual DNA concentrations within the pooled samples were present in equal concentration with the final concentration ranging from 1-25ng. Each reaction contained 1X PCR buffer (Qiagen, Valencia, CA), 200µM dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl₂, and 25pmol of the forward primer containing the universal primer sequence and the target specific sequence 5′-AGCGGATAACAATTTCACACAGGGAGCTAGCTTGGAAGAT

TGC-3' (SEQ ID NO: 41), 2pmol of the reverse primer 5'-GTCCAATATGCAAACAGTTG-3' (SEQ ID NO: 54), and 10pmol of a biotinylated universal primer complementary to the 5' end of the PCR amplicon BIO:5'-AGCGGATAACAATTTCACACAGG-3' (SEQ ID NO: 43).

After an initial round of amplification with the target with the specific forward and reverse primer, the 5' biotinylated universal primer can then be hybridized and acted as a forward primer thereby introducing a 5' biotin capture moiety into the molecule. The amplification protocol resulted in a 5'-biotinylated double stranded DNA amplicon and dramatically reduced the cost of high throughput genotyping by eliminating the need to 5' biotin label every forward primer used in a genotyping.

Themal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec; 72° C for 60 sec; 72° C 3min.

Immobilization of DNA

20

The 50 μ I PCR reaction was added to 25 μ L of streptavidin coated magnetic beads (Dynal, Oslo, Norway), which were prewashed three times and resuspended in 1M NH₄CI, 0.06M NH₄OH. The 5' end of one strand of the double stranded PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The hybridized but unbound strand was released from the double stranded amplicons by incubation in 100mM NaOH and washing of the beads three times with 10mM Tris pH 8.0.

Detection of AKAP10-5 using BiomassPROBE™ Assay

BiomassPROBE™ assay of primer extension analysis (see, U.S. Patent No. 6,043,031) of donor population for AKAP 10-5 (SEQ ID NO:

33) was performed. Genotyping using these methods was carried out by resuspending the DNA coated magnetic beads in 26mM Tris-HCL pH 9.5, 6.5 mM MgCl₂, 50mM dTTP, 50mM each of ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase (Ambersham), and 20pmol of a template specific oligonucleotide PROBE primer 5'-ACTGAGCCTGCTGCATAA-3' (SEQ ID NO: 44) (Operon). Primer extension occurs with three cycles of oligonucleotide primer with hybridization and extension. The extension products were analyzed after denaturation from the template with 50 mM NH₄Cl and transfer of 150 nL of each sample to a silicon chip preloaded with 150 nl of H3PA matrix material. The sample material was allowed to crystallize and analyzed by MALDI-TOF (Bruker, PerSeptive). The primer has a mass of 5483.6 daltons. The SNP results in the additional of a ddC to the primer, giving a mass of 5756.8 daltons for the extended product. The wild type results in the addition a T and ddG to the primer giving a mass of 6101 daltons.

The frequency of the SNP was measured in a population of age selected healthy individuals. Seven hundred thirteen (713) individuals under 40 years of age (360 females, 353 males) and 703 individuals over 60 years of age (322 females, 381 males) were tested for the presence of the SNP, AKAP10-5 (SEQ ID NO: 33). Results are presented below in Table 1.

AKAP10)-5 (2073V) fi		ABLE 1 comparison	in 2 age	groups
			<40	>60	delta G allele
Female	Alleles	*G	38.6	34.6	4.0
		*A	61.4	65.4	
	Genotypes	G	13.9	11.8	2.1
(T = 1)		GA	49.4	45.7	
``		Α	36.7	42.5	

25

20

15

25

				,		
	Male	Alleles	*G	41.4	37.0	4.4
			*A	58.6	63.0	
		Genotypes	G	18.4	10.8	7.7
5			GA	45.9	52.5	
-	÷		Α	35.7	36.7	-
	Total	Alleles	*G	40.0	35.9	4.1
			*A	60.0	64.1	·
10		Genotypes	G	16.1	11.2	4.9
			GA	47.7	49.4	
			Α	36.2	39.4	,

Figure 20 graphically shows these results of allele and genotype distribution in the age and sex stratified Caucasian population.

Morbidity marker 3: human methionine sulfoxide reductase A (msrA)

The age-related allele and genotype frequency of this marker in both genders and the entire population is shown in Figure 21. The decrease of the homozygous CC genotype in the older male population is highly significant.

Methionine sulfoxide reductase A (#63306)

PCR Amplification and BiomassPROBE assay detection of the human methioine sulfoxid reductase A (h-msr-A) in a healthy donor population PCR Amplification of donor population for h-msr-A

PCR primers were synthesized by OPERON using phosphoramidite chemistry. Amplification of the AKAP10 target sequence was carried out in single 50μ l PCR reaction with 100ng-1ug of pooled human genomic DNA templates in a 50μ l PCR reaction. Individual DNA concentrations within the pooled samples were present in an equal concentration with

the final concentration ranging from 1-25ng. Each reaction containing I X PCR buffer (Qiagen, Valencia, CA), $200\mu M$ dNTPs, 1U Hotstar Taq polymerase (Qiagen, Valencia, CA), 4mM MgCl₂, 25pmol of the forward primer containing the universal primer sequence and the target specific sequence 5'-TTTCTCTGCACAGAGAGGC-3' (SEQ ID NO: 49), 2pmol of the reverse primer 5'-AGCGGATAACAATTTCACACAGGGCTGAAATCCTTCGCTTTACC-3' (SEQ ID NO: 50), and 10pmol of a biotinylated universal primer complementary to the 5' end of the PCR amplicon 5'-AGCGGATAACAATTTCACACAGG-3' (SEQ ID NO: 51). After an initial 10 round of amplification of the target with the specific forward and reverse primers, the 5' biotinylated universal primer was then hybridized and acted as a reverse primer thereby introducing a 3' biotin capture moiety into the molecule. The amplification protocol results in a 5'-biotinylated 15 double stranded DNA amplicon and and dramatically reduces the cost of high throughput genotyping by eliminating the need to 5' biotin label each forward primer used in a genotyping. Thermal cycling was performed in 0.2mL tubes or 96 well plate using an MJ Research Thermal Cycler (calculated temperature) with the following cycling parameters: 94° C for 5 min; 45 cycles: 94° C for 20 sec, 56° C for 30 sec, 72° C for 60 sec; 20 72° C 3min.

Immobilization of DNA

The 50µl PCR reaction was added to 25ul of streptavidin coated magnetic bead (Dynal) prewashed three times and resuspended in 1M

NH₄Cl, 0.06M NH₄OH. The PCR amplicons were allowed to bind to the beads for 15 minutes at room temperature. The beads were then collected with a magnet and the supernatant containing unbound DNA was removed. The unbound strand was release from the double stranded

amplicons by incubation in 100mM NaOH and washing of the beads three times with 10mM Tris pH 8.0.

BiomassPROBE assay analysis of donor population for h-msr A

Genotyping using the BiomassPROBE assay methods was carried out by resuspending the he DNA coated magnetic beads in 26mM Tris-HCl pH 9.5, 6.5 mM MgCl₂, 50mM of dTTPs and 50mM each of ddCTP, ddATP, ddGTP, 2.5U of a thermostable DNA polymerase (Ambersham), and 20pmol of a template specific oligonucleotide PROBE primer 5'-CTGAAAAGGGAGAGAAAG-3' (Operon) (SEQ ID NO: 52).

Primer extension occurs with three cycles of oligonucleotide primer with hybridization and extension. The extension products were analyzed after denaturation from the template with 50mM NH₄Cl and transfer of 150nl each sample to a silicon chip preloaded with 150nl of H3PA matrix material. The sample material was allowed to crystallize and analyzed by MALDI-TOF (Bruker, PerSeptive). The SNP is represented as a T to C

MALDI-TOF (Bruker, PerSeptive). The SNP is represented as a T to C tranversion in the sequence of two ESTs. The wild type is represented by having a T at position 128 of GenBank Accession No. AW 195104, which represents the nucleotide sequence of an EST which is a portion of the wild type human msrA gene (SEQ ID NO: 39). The SNP is presented as a C at position 129 of GenBank Accession No. AW 874187, which represents the nucleotide sequence of an EST which is a portion of an allele of the human msrA gene (SEQ ID NO: 40).

In a genomic sequence the SNP is represented as an A to G transversion. The primer utilized in the BioMass probe reaction had a mass of 5654.8 daltons. In the presence of the SNP the primer is extended by the incorporation of a ddC and has a mass of 5928. In the presence of the wildtype the primer is extended by adding a dT and a DDC to produce a mass of 6232.1 daltons.

20

25

The frequency of the SNP was measured in a population of age selected healthy individuals. Five hundred fifty-two (552) individuals between the ages of 18-39 years (276 females, 276 males and 552 individuals between the age of 60-79 (184 females between the ages of 60-69, 368 males between the age of 60-79) were tested for the presence of the polymorphism localized in the nontranslated 3'region of h-msr-A.

Genotype difference between male age group among healthy individuals is significant. For the male population allele significance is p = 0.0009 and genotype significance is p = 0.003. The age-related allele and genotype frequency of this marker in both genders and the entire population is shown in Figure 21. The decrease of the homozygous CC genotype in the older male population is highly significant.

The polymorphism is localized in the non-translated 3'-region of the gene encoding the human methionine sulfoxide reductase (h-msrA). The exact localization is 451 base pairs downstream the stop codon (TAA). It is very likely that this SNP is in linkage disequilibrium (LD) with another polymorphism more upstream in the coding or promoter region; thus, it is not directly cause morbidity. The enzyme methionine sulfoxide reductase has been proposed to exhibit multiple biological functions. It may serve to repair oxidative protein damage but also play an important role in the regulation of proteins by activation or inactivation of their biological functions (Moskovitz et al. (1990) PNAS 95:14071-14075). It has also been shown that its activity is significantly reduced in brain tissues of Alzheimer patients (Gabbita et al., (1999) J. Neurochem 73:1660-1666). It is scientifically conceivable that proteins involved in the metabolism of reactive oxygen species are associated to disease.

20

CONCLUSION

The use of the healthy population provides for the identification of morbidity markers. The identification of proteins involved in the G-protein coupled signaling transduction pathway or in the detoxification of oxidative stress can be considered as convincing results. Further confirmation and validation of other potential polymorphisms already identified in *silico* in the gene encoding the human protein kinase A anchoring protein could even provide stronger association to morbidity and demonstrate that this gene product is a suitable pharmaceutical or diagnostic target.

EXAMPLE 4

MALDI-TOF Mass Spectrometry Analysis

All of the products of the enzyme assays listed below were analyzed by MALDI-TOF mass spectrometry. A diluted matrix solution (0.15 μ L) containing of 10:1 3-hydroxypicolinic acid:ammonium citrate in 1:1 water:acetonitrile diluted 2.5-fold with water was pipetted onto a SpectroChip (Sequenom, Inc.) and was allowed to crystallize. Then, 0.15 μ L of sample was added. A linear PerSeptive Voyager DE mass spectrometer or Bruker Biflex MALDI-TOF mass spectrometer, operating in positive ion mode, was used for the measurements. The sample plates were kept at 18.2 kV for 400 nm after each UV laser shot (approximate 250 laser shots total), and then the target voltage was raised to 20 kV. The original spectra were digitized at 500 MHz.

EXAMPLE 5

25 Sample Conditioning

Where indicated in the examples below, the products of the enzymatic digestions were purified with ZipTips (Millipore, Bedford, MA). The ZipTips were pre-wetted with 10 μ L 50% acetonitrile and equilibrated 4 times with 10 μ I 0.1 M TEAAc. The oligonucleotide fragments were-

25

bound to the C18 in the ZipTip material by continuous aspiration and dispension of each sample into the ZipTip. Each digested oligonucleotide was conditioned by washing with 10 μ L 0.1 M TEAAc, followed by 4 washing steps with 10 μ L H₂O. DNA fragments were eluted from the Ziptip with 7 μ L 50% acetonitrile.

Any method for condition the samples may be employed. Methods for conditioning, which generally is used to increase peak resolution, are well known (see, e.g., International PCT application No. WO 98/20019).

EXAMPLE 6

10 DNA Glycosylase-Mediated Sequence Analysis

DNA Glycosylases modifies DNA at each position that a specific nucleobase resides in the DNA, thereby producing abasic sites. In a subsequent reaction with another enzyme, a chemical, or heat, the phosphate backbone at each abasic site can be cleaved.

The glycosylase utilized in the following procedures was uracil-DNA glycosylase (UDG). Uracil bases were incorporated into DNA fragments in each position that a thymine base would normally occupy by amplifying a DNA target sequence in the presence of uracil. Each uracil substituted DNA amplicon was incubated with UDG, which cleaved each uracil base in the amplicon, and was then subjected to conditions that effected backbone cleavage at each abasic site, which produced DNA fragments. DNA fragments were subjected to MALDI-TOF mass spectrometry analysis. Genetic variability in the target DNA was then assessed by analyzing mass spectra.

Glycosylases specific for nucleotide analogs or modified nucleotides, as described herein, can be substituted for UDG in the following procedures. The glycosylase methods described hereafter, in conjunction with phosphate backbone cleavage and MALDI, can be used to analyze DNA fragments for the purposes of SNP scanning, bacteria

20

25

typing, methylation analysis, microsatellite analysis, genotyping, and nucleotide sequencing and re-sequencing.

A. Genotyping

A glycosylase procedure was used to genotype the DNA sequence encoding UCP-2 (Uncoupling Protein 2). The sequence for UCP-2 is deposited in GenBank under accession number AF096289. The sequence variation genotyped in the following procedure was a cytosine (C-allele) to thymine (T-allele) variation at nucleotide position 4790, which results in a alanine to valine mutation at position 55 in the UCP-2 polypeptide.

DNA was amplified using a PCR procedure with a 50 μ L reaction volume containing of 5 pmol biotinylated primer having the sequence 5′-TGCTTATCCCTGTAGCTACCCTGTCTTGGCCTTGCAGATCCAA-3′ (SEQ ID NO: 91), 15 pmol non-biotinylated primer having the sequence 5′-AGCGGATAACAATTTCACACAGGCCATCACACCGCGGTACTG-3′ (SEQ ID NO: 92), 200 μ M dATP, 200 μ M dCTP, 200 μ M dGTP, 600 μ M dUTP (to fully replace dTTP), 1.5 mM to 3 mM MgCl₂, 1 U of HotStarTaq polymerase, and 25 ng of CEPH DNA. Amplification was effected with 45 cycles at an annealing temperature of 56°C.

The amplification product was then immobilized onto a solid support by incubating 50 μ L of the amplification reaction with 5 μ L of prewashed Dynabeads for 20 minutes at room temperature. The supernatant was removed, and the beads were incubated with 50 μ L of 0.1 M NaOH for 5 minutes at room temperature to denature the double-stranded PCR product in such a fashion that single-stranded DNA was linked to the beads. The beads were then neutralized by three washes with 50 μ L 10 mM TrisHCl (pH 8). The beads were resuspended in 10 μ L of a 60mM TrisHCl/1mM EDTA (pH 7.9) solution, and 1 U uracil DNA glycosylase was added to the solution for 45 minutes at 37°C to remove uracil nucleotides present in the single-stranded DNA linked to the beads.

The beads were then washed two times with 25 μ L of 10 mM TrisHCl (pH 8) and once with 10 μ L of water. The biotinylated strands were then eluted from the beads with 12 μ L of 2 M NH₄OH at 60°C for 10 minutes. The backbone of the DNA was cleaved by incubating the samples for 10 min at 95°C (with a closed lid), and ammonia was evaporated from the samples by incubating the samples for 11 min at 80°C.

The cleavage fragments were then analyzed by MALDI-TOF mass spectrometry as described in Example 4. The T-allele generated a unique fragment of 3254 Daltons. The C-allele generated a unique fragment of 4788 Daltons. These fragements were distinguishable in mass spectra. Thus, the above-identified procedure was successfully utilized to genotype individuals heterozygous for the C-allele and T-allele in UCP-2.

B. Glycosylase Analysis Utilizing Pooled DNA Samples

The glycosylase assay was conducted using pooled samples to detect genetic variability at the UCP-2 locus. DNA of known genotype 15 was pooled from eleven individuals and was diluted to a fixed concentration of 5 $ng/\mu L$. The procedure provided in Example 3A was followed using 2 pmol of forward primer having a sequence of 5'-CCCAGTCACGACGTTGTAAAACGTCTTGGCCTTGCAGATCCAAG- 3' (SEQ ID NO: 93) and 15 pmol of reverse primer having the sequence 5'-20 AGCGGATAACAATTTCACACAGGCCATCACACCGCGGTACTG-3' (SEQ ID NO: 94). In addition, 5 pmol of biotinylated primer having the sequence 5'bioCCCAGTCACGACGTTGTAAAACG 3' (SEQ ID NO: 97) may be introduced to the PCR reaction after about two cycles. The fragments were analyzed via MALDI-TOF mass spectroscopy (Example 4). As determined in Example 3A, the T-allele, which generated a unique fragment of 3254 Daltons, could be distinguished in mass spectra from the C-allele, which generated a unique fragment of 4788 Daltons. Allelic frequency in the pooled samples was quantified by integrating the area

25

under each signal corresponding to an allelic fragment. Integration was accomplished by hand calculations using equations well known to those skilled in the art. In the pool of eleven samples, this procedure suggested that 40.9% of the individuals harbored the T allele and 59.09% of the individuals harbored the C allele.

C. Glycosylase-Mediated Microsatellite Analysis

A glycosylase procedure was utilized to identify microsatellites of the Bradykinin Receptor 2 (BKR-2) sequence. The sequence for BKR-2 is deposited in GenBank under accession number X86173. BKR-2 includes a SNP in the promoter region, which is a C to T variation, as well as a SNP in a repeated unit, which is a G to T variation. The procedure provided in Example 3A was utilized to identify the SNP in the promotor region, the SNP in the microsattelite repeat region, and the number of repeated units in the microsattelite region of BKR-2. Specifically, a forward PCR primer having the sequence 5'-CTCCAGCTGGGCAGGAGTGC-3' (SEQ ID NO: 95) and a reverse primer having the sequence 5'-CACTTCAGTCGCTCCCT-3' (SEQ ID NO: 96) were utilized to amplify BKR-2 DNA in the presence of uracil. The amplicon was fragmented by UDG followed by backbone cleavage. The cleavage fragments were analyzed by MALDI-TOF mass spectrometry as described in Example 4.

With regard to the SNP in the BKR-2 promotor region having a C to T variation, the C-allele generated a unique fragment having a mass of 7342.4 Daltons and the T-allele generated a unique fragment having a mass of 7053.2 Daltons. These fragments were distinguishable in mass spectra. Thus, the above-identified procedure was successfully utilized to genotype individuals heterozygous for the C-allele and T-allele in the promotor region of BKR-2.

With regard to the SNP in the BKR-2 repeat region having a G to T variation, the T-allele generated a unique fragment having a mass of 1784 Daltons, which was readily detected in a mass spectrum. Hence, the presence of the T-allele was indicative of the G to T sequence variation in the repeat region of BKR-2.

In addition, the number of repeat regions was distinguished between individuals having two repeat sequences and individuals having three repeat sequences in BKR-2. The DNA of these individuals did not harbor the G to T sequence variation in the repeat sequence as each repeat sequence contained a G at the SNP locus. The number of repeat regions was determined in individual samples by calculating the area under a signal corresponding to a unique DNA fragment having a mass of 2771.6 Daltons. This signal in spectra generated from individuals having two repeat regions had an area that was thirty-three percent less than the area under the same signal in spectra generated from individuals having three repeat regions. Thus, the procedures discussed above can be utilized to genotype individuals for the number of repeat sequences present in BKR-2.

D. Bisulfite Treatment Coupled with Glycosylase Digestion

Bisulfite treatment of genomic DNA can be utilized to analyze positions of methylated cytosine residues within the DNA. Treating nucleic acids with bisulfite deaminates cytosine residues to uracil residues, while methylated cytosine remains unmodified. Thus, by comparing the sequence of a PCR product generated from genomic DNA that is not treated with bisulfite with the sequence of a PCR product generated from genomic DNA that is treated with bisulfite, the degree of methylation in a nucleic acid as well as the positions where cytosine is methylated can be deduced.

25

Genomic DNA (2 μ g) was digested by incubation with 1 μ L of a restriction enzyme at 37°C for 2 hours. An aliquot of 3 M NaOH was added to yield a final concentration of 0.3M NaOH in the digestion solution. The reaction was incubated at 37°C for 15 minutes followed by treatment with 5.35M urea, 4.44M bisulfite, and 10mM hydroquinone, where the final concentration of hydroquinone is 0.5 mM.

The sample that was treated with bisulfite (sample A) was compared to the same digestion sample that had not undergone bisulfite treatment (sample B). After sample A was treated with bisulfite as described above, sample A and sample B were amplified by a standard. PCR procedure. The PCR procedure included the step of overlaying each sample with mineral oil and then subjecting the sample to thermocycling (20 cycles of 15 minutes at 55°C followed by 30 seconds at 95°C). The PCR reaction contained four nucleotide bases, C, A, G, and U. The mineral oil was removed from each sample, and the PCR products were purified with glassmilk. Sodium iodide (3 volumes) and glassmilk (5 μ L) were added to samples A and B. The samples were then placed on ice for 8 minutes, washed with 420 μ L cold buffer, centrifuged for 10 seconds, and the supernatant fractions were removed. This process was repeated twice and then 25 μL of water was added. Samples were incubated for 5 minutes at 37 °C, were centrifuged for 20 seconds, and the supernatant fraction was collected, and then this incubation/centrifugation/supernatant fraction collection procedure was repeated. 50 μ L 0.1 M NaOH was then added to the samples to denature the DNA. The samples were incubated at room temperature for 5 minutes, washed three times with 50 μ L of 10 mM TrisHCl (pH 8), and resuspended in 10 μ L 60mM TrisHCI/1mM EDTA, pH 7.9.

The sequence of PCR products from sample A and sample B were then treated with 2U of UDG (MBI Fermentas) and then subjected to

25

backbone cleavage, as described herein. The resulting fragments from each of sample A and sample B were analyzed by MALDI-TOF mass spectroscopy as described in Example 4. Sample A gave rise to a greater number of fragments than the number of fragments arising from sample B, indicative that the nucleic acid harbored at least one methylated cytosine moiety.

EXAMPLE 7

Fen-Ligase-Mediated Haplotyping

Haplotyping procedures permit the selection of a fragment from one of an individual's two homologous chromosomes and to genotype linked SNPs on that fragment. The direct resolution of haplotypes can yield increased information content, improving the diagnosis of any linked disease genes or identifying linkages associated with those diseases. In previous studies, haplotypes were typically reconstructed indirectly through pedigree analysis (in cases where pedigrees were available) through laborious and unreliable allele-specific PCR or through single-molecule dilution methods well known in the art.

A haplotyping procedure was used to determine the presence of two SNPs, referred to as SNP1 and SNP2, located on one strand in a DNA sample. The haplotyping procedure used in this assay utilized Fen-1, a site-specific "flap" endonuclease that cleaves DNA "flaps" created by the overlap of two oligonucleotides hybridized to a target DNA strand. The two overlapping oligonucleotides in this example were short arm and long arm allele-specific adaptors. The target DNA was an amplified nucleic acid that had been denatured and contained SNP1 and SNP2.

The short arm adaptor included a unique sequence not found in the target DNA. The 3' distal nucleotide of the short arm adaptor was identical to one of the SNP1 alleles. Moreover, the long arm adaptor included two regions: a 3' region complementary to the short arm and-a

25

5'gene-specific region complementary to the fragment of interest adjacent to the SNP. If there was a match between the adaptor and one of the homologues, the Fen enzyme recognized and cleaved the overlapping flap. The short arm of the adaptor was then ligated to the remainder of the target fragment (minus the SNP site). This ligated fragment was used as the forward primer for a second PCR reaction in which only the ligated homologue was amplified. The second PCR product (PCR2) was then analyzed by mass spectrometry. If there was no match between the adaptors and the target DNA, there was no overlap, no cleavage by Fen-1, and thus no PCR2 product of interest.

If there was more than one SNP in the sequence of interest, the second SNP (SNP2) was found by using an adaptor that was specific for SNP2 and hybridizing the adaptor to the PCR2 product containing the first SNP. The Fen-ligase and amplification procedures were repeated for the PCR2 product containing the first SNP. If the amplified product yielded a second SNP, then SNP1 and SNP2 were on the same fragment.

If the SNP is unknown, then four allele-specific adaptors (e.g. C, G, A, and T) can be used to hybridize with the target DNA. The substrates are then treated with the Fen-ligase protocol, including amplification. The PCR2 products may be analyzed by PROBE, as described herein, to determine which adaptors were hybridized to the DNA target and thus identify the SNPs in the sequence.

A Fen-ligase assay was used to detect two SNPs present in Factor VII. These SNPs are located 814 base pairs apart from each other. SNP1 was located at position 8401 (C to T), and SNP2 was located at 9215 (G to A) (SEQ ID #).

A. First Amplification Step

A PCR product (PCR1) was generated for a known heterozygous individual at SNP1, a short distance from the 5' end of the SNP.

Specifically, a 10 μ L PCR reaction was performed by mixing 1.5 mM MgCl₂, 200 μ M of each dNTP, 0.5 U HotStar polymerase, 0.1 μ M of a forward primer having the sequence 5'-GCG CTC CTG TCG GTG CCA (SEQ ID NO: 56), 0.1 μ M of a reverse primer having the sequence 5'-GCC TGA CTG GTG GGG CCC (SEQ ID NO: 57), and 1 ng of genomic DNA. The annealing temperature was 58°C, and the amplification process yielded fragments that were 861 bp in length.

The PCR1 reaction mixture was divided in half and was treated with an exonuclease 1/SAP mixture (0.22 μ L mixture/5 μ L PCR1 reaction) which contained 1.0 μ L SAP and 0.1 μ L exon1. The exonuclease treatment was done for 30 minutes at 37°C and then 20 minutes at 85°C to denature the DNA.

B. Adaptor Oligonucleotides

A solution of allele-specific adaptors (C and T), containing of one long and one short oligonucleotide per adaptor, was prepared. The long arm and short arm oligonucleotides of each adaptor (10µM) were mixed in a 1:1 ratio and heated for 30 seconds at 95°C. The temperature was reduced in 2°C increments to 37°C for annealing. The C-adaptor had a short arm sequence of 5'-CAT GCA TGC ACG GTC (SEQ ID NO: 58) and a long arm sequence of 5'-CAG AGA GTA CCC CTC GAC CGT GCA TGC ATG (SEQ ID NO: 59). Hence, the long arm of the adaptor was 30 bp (15 bp gene-specific), and the short arm was 15bp. The T-adaptor had a short arm sequence of 5'-CAT GCA TGC ACG GTT (SEQ ID NO: 60) and a long arm sequence of 5'-GTA CGT ACG TGC CAA CTC CCC ATG AGA GAC (SEQ ID NO: 61). The adaptor could also have a hairpin structure in which the short and long arm are separated by a loop containing of 3 to 10 nucleotides (SEQ ID NO: 118).

C. FEN-ligase reaction

In two tubes (one tube for each allele-specific adaptor per sample) was placed a solution (Solution A) containing of 3.5 μ l 10 mM 16%PEG/50 mM MOPS, 1.2 μ l 25 mM MgCl₂, 1.5 μ l 10X Ampligase Buffer, and 2.5 μ l PCR1. Each tube containing Solution A was incubated at 95°C for 5 minutes to denature the PCR1 product. A second solution (Solution B) containing of 1.65 μ l Ampligase (Thermostable ligase, Epicentre Technologies), 1.65 μ l 200ng/ μ l MFEN (from *Methanocuccus jannaschii*), and 3.0 μ l of an allele specific adaptor (C or T) was prepared. Thus, different variations of Solution B, each variation containing of different allele-specific adaptors, were made. Solution B was added to Solution A at 95°C and incubated at 55°C for 3 hours. The total reaction volume was 15.0 μ l per adaptor-specific reaction. For a bi-allelic system, 2 x 15.0 μ l reactions were required.

The Fen-ligase reaction in each tube was then deactivated by adding 8.0 μ l 10 mM EDTA. Then, 1.0 μ l exoIII/Buffer (70%/30%) solution was added to each sample and incubated 30 minutes at 37°C, 20 minutes at 70°C (to deactivate exoIII), and 5 minutes at 95°C (to denature the sample and dissociate unused adaptor from template). The samples were cooled in an ice slurry and purified on UltraClean PCR Clean-up (MoBio) spin columns which removed all fragments less than 100 base pairs in length. The fragments were eluted with 50 μ l H₂0.

D. Second Amplification Step

20

25

A second amplification reaction (PCR2) was conducted in each sample tube using the short arm adaptor (C or T) sequence as the forward primer (minus the SNP1 site). Only the ligated homologue was amplified. A standard PCR reaction was conducted with a total volume of 10.0 μ l containing of 1X Buffer (final concentration), 1.5 mM final concentration MgCl₂, 200 μ M final concentration dNTPs, 0.5 U HotStar polymerase, 0.1 μ M final concentration forward primer 5'-CAT GCA TGC ACG GT (SEQ-ID

15

NO: 62), 0.1 μ M final concentration reverse primer 5'-GCC TGA CTG GTG GGG CCC (SEQ ID NO: 63), and 1.0 μ l of the purified FEN-ligase reaction solution. The annealing temperature was 58°C. The PCR2 product was analyzed by MALDI TOF mass spectroscopy as described in Example 4. The mass spectrum of Fen SNP1 showed a mass of 6084.08 Daltons, representing the C allele.

E. Genotyping Additional SNPs

The second SNP (SNP2) can be found by using an adaptor that is specific for SNP2 and hybridizing that adaptor to the PCR2 product containing the first SNP. The Fen-ligase and amplification procedures are repeated for the PCR2 product containing the first SNP. If the amplified product yields a second SNP, then SN1 and SN2 are on the same fragment. The mass spectrum of SNP2, representing the T allele, showed a mass of 6359.88 Daltons.

This assay can also be performed upon pooled DNA to yield haplotype frequencies as described herein. The Fen-ligase assay can be used to analyze multiplexes as described herein.

EXAMPLE 8

Nickase-Mediated Sequence Analysis

20 A DNA nickase, or DNase, was used to recognize and cleave one strand of a DNA duplex. Two nickases usd were NY2A nickase and NYS1 nickase (Megabase) which cleave DNA at the following sites:

NY2A: 5'...R AG...3'

 $3'...Y \downarrow TC...5'$ where R = A or G and Y = C or T

25 NYS1: 5'...↓CC[A/G/T]...3'

3'... GG[T/C/A]...5'.

A. Nickase Digestion

Tris-HCI (10 mM), KCI (10 mM, pH 8.3), magnesium acetate (25 mM), BSA (1 mg/mL), and 6 U of Cvi NY2A or Cvi NYS1 Nickase (Megabase Research) were added to 25 pmol of double-stranded oligonucleotide template having a sequence of 5'-CGC AGG GTT TCC TCG TCG CAC TGG GCA TGT G-3' (SEQ ID NO: 90, Operon, Alameda, CA) synthesized using standard phosphoramidite chemistry. With a total volume of 20µL, the reaction mixture was incubated at 37°C for 5 hours, and the digestion products were purified using ZipTips (Millipore, Bedford, 10 MA) as described in Example 5. The samples were analyzed by MALTY-TOM mass spectroscopy as described in Example 1. The nickase Cvi NY2A yielded three fragments with masses 4049.76 Daltons, 5473.14 Daltons, and 9540.71 Daltons. The Cvi NYS1 nickase yielded fragments with masses 2063.18 Daltons, 3056.48 Daltons, 6492.81 Daltons, and 7450.14 Daltons.

B. Nickase Digestion of Pooled Samples

DQA (HLA ClassII-DQ Alpha, expected fragment size = 225bp) was amplified from the genomic DNA of 100 healthy individuals. DQA was amplified using standard PCR chemistry in a reaction having a total volume of 50 μL containing of 10 mM Tris-HCl, 10 mM KCl (pH 8.3), 2.5 mM MgCl₂, 200 μM of each dNTP, 10 pmol of a forward primer having the sequence 5′-GTG CTG CAG GTG TAA ACT TGT ACC AG-3′(SEQ ID NO: 64), 10 pmol of a reverse primer having the sequence 5′-CAC GGA TCC GGT AGC AGC GGT AGA GTT G-3′(SEQ ID NO: 65), 1 U DNA polymerase (Stoffel fragment, Perkin Elme r), and 200ng human genomic DNA (2ng DNA/individual). The template was denatured at 94°C for 5 minutes. Thermal cycling was continued with a touch-down program that included 45 cycles of 20 seconds at 94°C, 30 seconds at 56°C, 1

25

minute at 72°C, and a final extension of 3 minutes at 72°C. The crude PCR product was used in the subsequent nickase reaction.

The unpurified PCR product was subjected to nickase digestion. Tris-HCl (10 mM), KCl (10 mM, pH 8.3), magnesium acetate (25mM), BSA (1 mg/mL), and 5 U of Cvi NY2A or Cvi NYS1 Nickase (Megabase Research) were added to 25 pmol of the amplified template with a total reaction volume of 20µL. The mixture was then incubated at 37°C for 5 hours. The digestion products were purified with either ZipTips (Millipore, Bedford, MA) as described in Example 5. The samples were analyzed by MALDI-TOF mass spectroscopy as described in Example 4. This assay can also be used to do multiplexing and standardless genotyping as described herein.

To simplify the nickase mass spectrum, the two complementary strands can be separated after digestion by using a single-stranded undigested PCR product as a capture probe. This probe (preparation shown below in Example 8C) can be hybridized to the nickase fragments in hybridization buffer containing 200 mM sodium citrate and 1% blocking reagent (Boehringer Mannheim). The reaction is heated to 95°C for 5 minutes and cooled to room temperature over 30 minutes by using a thermal cycler (PTC-200 DNA engine, MJ Research, Waltham, MA). The capture probe-nickase fragment is immobilized on 140 μ g of streptavidin-coated magnetic beads. The beads are subsequently washed three times with 70 mM ammonium citrate. The captured single-stranded nickase fragments are eluted by heating to 80°C for 5 minutes in 5 μ L of 50 mM ammonium hydroxide.

C. Preparation of Capture Probe

The capture probe is prepared by amplifying the human β -globin gene (3' end of intron 1 to 5' end of exon 2) via PCR methods in a total volume of 50 μ L containing of GeneAmp 1XPCR Buffer II, 10 mM Tris-

HCI, pH 8.3, 50 mM KCI, 2 mM MgCI₂, 0.2 mM dNTP mix, 10pmol of each primer (forward primer 5'-ACTGGGCATGTGGAGACAG-3'(SEQ ID NO: 66) and biotinylated reverse primer bio5'-GCACTTTCTTGCCATGAG-3'(SEQ ID: 67), 2 U of AmpliTaq Gold, and 200 ng of human genomic DNA. The template is denatured at 94°C for 8 minutes. Thermal cycling is continued with a touch-down program that included 11 cycles of 20 seconds at 94°C, 30 seconds at 64°C, 1 minute at 72°C; and a final extension of 5 minutes at 72°C. The amplicon is purified using UltraClean PCR clean-up kit (MO Bio Laboratories, Solano Beach, CA).

10

15

20

25

EXAMPLE 9

Multiplex Type IIS SNP Assay

A Type IIS assay was used to identify human gene sequences with known SNPs. The Type IIS enzyme used in this assay was Fok I which effected double-stranded cleavage of the target DNA. The assay involved the steps of amplification and Fok I treatment of the amplicon. In the amplification step, the primers were designed so that each PCR product of a designated gene target was less than 100 bases such that a Fok I recognition sequence was incorporated at the 5' and 3' end of the amplicon. Therefore, the fragments that were cleaved by Fok I included a center fragment containing the SNP of interest.

Ten human gene targets with known SNPs were analyzed by this assay. Sequences of the ten gene targets, as well as the primers used to amplify the target regions, are found in Table 5. The ten targets were lipoprotein lipase, prothrombin, factor V, cholesterol ester transfer protein (CETP), factor VII, factor XIII, HLA-H exon 2, HLA-H exon 4, methylenetetrahydrofolate reductase (MTHR), and P53 exon 4 codon 72.

Amplification of the ten human gene sequences were carried out in a single 50 μ L volume PCR reaction with 20 ng of human genomic DNA

template in 5 PCR reaction tubes. Each reaction vial contained 1X PCR buffer (Qiagen), 200 μ M dNTPs, 1U Hotstar Taq polymerase (Qiagen), 4 mM MgCl₂, and 10pmol of each primer. US8, having sequence of 5'TCAGTCACGACGTT3'(SEQ ID NO: 68), and US9, having sequence of 5'CGGATAACAATTTC3'(SEQ ID NO: 69), were used for the forward and reverse primers respectively. Moreover, the primers were designed such that a Fok I recognition site was incorporated at the 5' and 3' ends of the amplicon. Thermal cycling was performed in 0.2 mL tubes or a 96 well plate using a MJ Research Thermal Cycler (calculated temperature) with the following cycling parameters: 94°C for 5 minutes; 45 cycles: 94°C for 20 seconds, 56°C for 20 seconds, 72°C for 60 seconds; and 72°C for 3 minutes.

Following PCR, the sample was treated with 0.2 U Exonuclease I (Amersham Pharmacia) and S Alkaline Phosphotase (Amersham Pharmacia) to remove the unincorporated primers and dNTPs. Typically, 0.2 U of exonuclease I and SAP were added to 5 μ L of the PCR sample. The sample was then incubated at 37°C for 15 minutes. Exonuclease I and SAP were then inactivated by heating the sample up to 85°C for 15 minutes. Fok I digestion was performed by adding 2 U of Fok I (New England Biolab) to the 5 uL PCR sample and incubating at 37°C for 30 minutes. Since the Fok I restriction sites are located on both sides of the amplicon, the 5' and 3' cutoff fragments have higher masses than the center fragment containing the SNP. The sample was then purified by anion exchange and analyzed by MALDI-TOF mass spectrometry as 25 described in Example 4. The masses of the gene fragments from this multiplexing experiment are listed in Table 6. These gene fragments were resolved in mass spectra thereby allowing multiplex analysis of sequence variability in these genes.

> Table 5 Genes for Multiplex Type IIS Assay

20

Gene	Sequence	Seq. ID No.	Primers	Se ID
Lipoprotein Lipase (Asn291Ser)	cctttgagaa agggctctgc ttgagttgta gaaagaaccg ctgcaacaat ctgggctatg agatca(a+g)taa agtcagagcc aaaagaagca gcaaaatgta	98-99	5' caattcatcgctggatgcaatct gggctatgagatc 3' 5' caatttcacacagcggatgcttct tttggctctgact 3'	<u>70</u>
Prothrombin	26731 gaattatttt tgtgitteta aaactatggt te <u>ccaataaa aqtqaetete</u> 26781 <u>agelg>alageete aatgeteeca</u> gtgetattea tgggeagete tetgggetea	100- 101	5' tcagtcacgacgttggat <u>gcact</u> tagataacaatttcggat <u>gcact</u> 5' cggataacaatttcggat <u>gccca</u>	73
Factor V (Arg506Gln)	taataggact actictaatc tgtaa <u>qaqca</u> gatccctgga caggc[g> a]agga atacaggtat tttgtccttg aagtaacctt tcag	102- 103	5' tcagtcacgacgttggatgagca gatccctggacaggc 3' 5' cggataacaatttcggatggaca aatacctgtattcc 3'	74
Cholesterol ester transfer protein (CETP) (I405V)	1261 ctcaccatgg gcatttgatt <u>gcagagcage</u> tccgagtcc[g > a] tccagagctt 1311 cctgcagtca atgatcaccg ctgtgggcat ccctgaggtc atgtctcgta	104- 105	5' tcagtcacgacgttggatgcaga gcagetccgagtc 3' 5' cageggtgatcattggatgcagg aagetctgg 3'	7(
Factor VII (R353Q)	1221 agcaeggact cetgeaaggg ggacagtgga ggc <u>ccacatq ccacccacta</u> 1271 <u>c</u> cla>glgggca <u>cq tqqtacetqa</u> <u>cqqqcatcgt</u> cagetggggc cagggctgcg	106- 107	5' tcagtcacgacgttggatgccca catgccaccactac 3' 5' cggataacaatttcggatgcccg tcaggtaccacg 3'	7:
Factor XIII (V,34L)	111 caataactot aatgoagogg aagatgaco <u>t</u> gcccacagtg gagottoagg 161 gc[g+t]tggtgcc ccgggg <u>cqtc</u> aacetgcaag gtatgagcat accccctto	108- 109	5' tcagtcacaccttgcaggatgacg cagtggagcttcag 3' 5' gctcataccttgcaggatgacg 3'	8
HLA-H exon 2 (His63Asp)	361 ttgaagetttgggetaegtg gatgaeeage tgttegtgttetatgat(e+g)at 411 gagagtegee gtgtggagee eegaacteea tgggttteea gtagaattte	110- 111	5' tcagtcacgacgttgggatgacca getgttcgtgttc3' 5' tacatggagttcggggatgacca cggcactctc 3'	8
HLA-H exon 4 (Cys282Tyr)	1021 ggataacett ggetgtaece ce <u>tggggaag</u> agcagagata taegt grajecag	112-	5' tcagtcacgacgttggatgggga agagcagagatatacgt 3'	8
`	attgtgatet gggageeete		5' gagggetgatccaggatgggt getccae 3'	1.

Gene	Sequence	Seq. ID No.	Primers	Seq. ID No.
Methylentetrahy drofolateredctas e (MTHR) (Ala222Val)	761 tgeagcactt geegge geeg gtgtctgcgg geglc>tlcgatttcatcatcacg 811 cagcttttctttgaggctga cacattcttc	114- 115	5' gaggagetgatecaggatgagt agagcagagatatacgt 3' 5' gctcacgacgttggatgagagagagagagagagagagaga	86
P53 Exon4 Codon 72 (Arg72Pro)	12101 tccagatgaa geteccagaa tgccagagge tgetecce(g>e)c gtggeccetg 12151 <u>caccagcage tcctacaecg</u> geggeccetg	116- 117	5' geogeoggtgtaggatgetgetg agge 3' 5' gtge 3'	88 89

Table 6
The mass of Center Fragments for Ten Different SNP Typing by IIS Assay

Gene	Tbr (_{ve}	LPL(Am291Ser)	Prothr	Prothrombin FV(Ar9506 ^{9h})	FV(Are	506 ^{gh})	CETP(CETP('405')	FVII(*353°)	353°)	FXIII	FXIII(^v 34)
Genotype A	۷	0,	ŋ	A	ŋ	٨	ပ	4	g	٨	ģ	۳
+ strand 6213 6229 5845 5829 5677 5661 3388 3372 6128 6112 5058 5033 (Da)	6213	6229	5845	5829	5677	5661	3388	3372	6128	6112	5058	5033
strand mass (Da)	6129	6129 6114 5949 5964 5472 5487 3437 3452 6174 6189 4916 4940	5949	5964	5472	5487	3437	3452	6174	6189	4916	4940

Gene	Ŧ	Hlah2	₩	Hlah4	MTHR(MTHR(**222**)	P53exor	P53exon4(Am72Pm)
Genotype	O .	9	<u>ຶ່</u> ງ	A	U	-	ŋ	ပ
+ strand mass (Da)	5889	5929	4392	4392 4376 4400	4400	4415	4586	4546
- strand mass - (Da)	5836	5836 5796 4319 4334	4319	4334	4368	4352	4724	4764

-90-

EXAMPLE 10

Exemplary use of parental medical history parameter for stratification of healthy datebase

A healthy database can be used to associate a disease state with a specific allele (SNP) that has been found to show a strong association between age and the allele, in particular the homozygous genotype. The method involves using the same healthy database used to identify the age dependent association, however stratification is by information given by the donors about common disorders from which their parents suffered (the donor's familial history of disease). There are three possible answers a donor could give about the health status of their parents: neither were affected, one was affected or both were affected. Only donors above a certain minimum age, depending on the disease, are utilized, as the donors parents must be old enough to to have exhibited clinical disease phenotypes. The genotype frequency in each of these groups is determined and compared with each other. If there is an association of the marker in the donor to a disease the frequency of the heterozyous genotype will be increased. The frequency of the homozygous genotype should not increase, as it should be significantly underrepresented in the healthy population.

25

30

-91-

EXAMPLE 11

Method and Device for Identifying a Biological Sample Description

In accordance with the present invention, a method and device for identifying a biological sample is provided. Referring now to FIG. 24, an apparatus 10 for identifying a biological sample is disclosed. The apparatus 10 for identifying a biological sample generally comprises a mass spectrometer 15 communicating with a computing device 20. In a preferred embodiment, the mass spectrometer may be a MALDI-TOF mass spectrometer manufactured by Bruker-Franzen Analytik GmbH; however, it will be appreciated that other mass spectrometers can be substituted. The computing device 20 is preferably a general purpose computing device. However, it will be appreciated that the computing device could be alternatively configured, for example, it may be integrated with the mass spectrometer or could be part of a computer in a larger network system.

The apparatus 10 for identifying a biological sample may operate as an automated identification system having a robot 25 with a robotic arm 27 configured to deliver a sample plate 29 into a receiving area 31 of the mass spectrometer 15. In such a manner, the sample to be identified may be placed on the plate 29 and automatically received into the mass spectrometer 15. The biological sample is then processed in the mass spectrometer to generate data indicative of the mass of DNA fragments in the biological sample. This data may be sent directly to computing device 20, or may have some preprocessing or filtering performed within the mass spectrometer. In a preferred embodiment, the mass spectrometer 15 transmits unprocessed and unfiltered mass spectrometry data to the computing device 20. However, it will be appreciated that the analysis in the computing device may be adjusted to accommodate preprocessing or filtering performed within the mass spectrometer.

Referring now to FIG. 25, a general method 35 for identifying a biological sample is shown. In method 35, data is received into a computing device from a test instrument in block 40. Preferably the data is received in a raw, unprocessed and unfiltered form, but alternatively may have some form of

15

20

25

filtering or processing applied. The test instrument of a preferred embodiment is a mass spectrometer as described above. However, it will be appreciated that other test instruments could be substituted for the mass spectrometer.

The data generated by the test instrument, and in particular the mass spectrometer, includes information indicative of the identification of the biological sample. More specifically, the data is indicative of the DNA composition of the biological sample. Typically, mass spectrometry data gathered from DNA samples obtained from DNA amplification techniques are noisier than, for example, those from typical protein samples. This is due in part because protein samples are more readily prepared in more abundance, and protein samples are more easily ionizable as compared to DNA samples. Accordingly, conventional mass spectrometer data analysis techniques are generally ineffective for DNA analysis of a biological sample. To improve the analysis capability so that DNA composition data can be more readily discerned, a preferred embodiment uses wavelet technology for analyzing the DNA mass spectrometry data. Wavelets are an analytical tool for signal processing, numerical analysis, and mathematical modeling. Wavelet technology provides a basic expansion function which is applied to a data set. Using wavelet decomposition, the data set can be simultaneously analyzed in the time and frequency domains. Wavelet transformation is the technique of choice in the analysis of data that exhibit complicated time (mass) and frequency domain information, such as MALDI-TOF DNA data. Wavelet transforms as described herein have superior denoising properties as compared to conventional Fourier analysis techniques. Wavelet transformation has proven to be particularly effective in interpreting the inherently noisy MALDI-TOF spectra of DNA samples. In using wavelets, a "small wave" or "scaling function" is used to transform a data set into stages, with each stage representing a frequency component in the data set. Using wavelet transformation, mass spectrometry data can be processed, filtered, and analyzed with sufficient discrimination to be useful for identification of the DNA composition for a biological sample.

Referring again to FIG. 25, the data received in block 40 is denoised in block 45. The denoised data then has a baseline correction applied in block 50.

15

20

25

30

A baseline correction is generally necessary as data coming from the test instrument, in particular a mass spectrometer instrument, has data arranged in a generally exponentially decaying manner. This generally exponential decaying arrangement is not due to the composition of the biological sample, but is a result of the physical properties and characteristics of the test instrument, and other chemicals involved in DNA sample preparation. Accordingly, baseline correction substantially corrects the data to remove a component of the data attributable to the test system, and sample preparation characteristics.

After denoising in block 45 and the baseline correction in block 50, a signal remains which is generally indicative of the composition of the biological sample. However, due to the extraordinary discrimination required for analyzing the DNA composition of the biological sample, the composition is not readily apparent from the denoised and corrected signal. For example, although the signal may include peak areas, it is not yet clear whether these "putative" peaks actually represent a DNA composition, or whether the putative peaks are result of a systemic or chemical aberration. Further, any call of the composition of the biological sample would have a probability of error which would be unacceptable for clinical or therapeutic purposes. In such critical situations, there needs to be a high degree of certainty that any call or identification of the sample is accurate. Therefore, additional data processing and interpretation is necessary before the sample can be accurately and confidently identified.

Since the quantity of data resulting from each mass spectrometry test is typically thousands of data points, and an automated system may be set to perform hundreds or even thousands of tests per hour, the quantity of mass spectrometry data generated is enormous. To facilitate efficient transmission and storage of the mass spectrometry data, block 55 shows that the denoised and baseline corrected data is compressed.

In a preferred embodiment, the biological sample is selected and processed to have only a limited range of possible compositions. Accordingly, it is therefore known where peaks indicating composition should be located, if present. Taking advantage of knowing the location of these expected peaks, in block 60 the method 35 matches putative peaks in the processed signal to the

20

25

30

location of the expected peaks. In such a manner, the probability of each putative peak in the data being an actual peak indicative of the composition of the biological sample can be determined. Once the probability of each peak is determined in block 60, then in block 65 the method 35 statistically determines the composition of the biological sample, and determines if confidence is high enough to calling a genotype.

Referring again to block 40, data is received from the test instrument, which is preferably a mass spectrometer. In a specific illustration, FIG. 26 shows an example of data from a mass spectrometer. The mass spectrometer data 70 generally comprises data points distributed along an x-axis 71 and a y-axis 72. The x-axis 71 represents the mass of particles detected, while the y-axis 72 represents a numerical concentration of the particles. As can be seen in FIG. 26, the mass spectrometry data 70 is generally exponentially decaying with data at the left end of the x-axis 73 generally decaying in an exponential manner toward data at the heavier end 74 of the x-axis 71. However, the general exponential presentation of the data is not indicative of the composition of the biological sample, but is more reflective of systematic error and characteristics. Further, as described above and illustrated in FIG. 26, considerable noise exists in the mass spectrometry DNA data 70.

Referring again to block 45, where the raw data received in block 40 is denoised, the denoising process will be described in more detail. As illustrated in FIG. 25, the denoising process generally entails 1) performing a wavelet transformation on the raw data to decompose the raw data into wavelet stage coefficients; 2) generating a noise profile from the highest stage of wavelet coefficients; and 3) applying a scaled noise profile to other stages in the wavelet transformation. Each step of the denoising process is further described below.

Referring now to FIG. 27, the wavelet transformation of the raw mass spectrometry data is generally diagramed. Using wavelet transformation techniques, the mass spectrometry data 70 is sequentially transformed into stages. In each stage the data is represented in a high stage and a low stage, with the low stage acting as the input to the next sequential stage. For example, the mass spectrometry data 70 is transformed into stage 0 high data

25

30

82 and stage 0 low data 83. The stage 0 low data 83 is then used as an input to the next level transformation to generate stage 1 high data 84 and stage 1 low data 85. In a similar manner, the stage 1 low data 85 is used as an input to be transformed into stage 2 high data 86 and stage 2 low data 87. The transformation is continued until no more useful information can be derived by further wavelet transformation. For example, in the preferred embodiment a 24-point wavelet is used. More particularly a wavelet commonly referred to as the Daubechies 24 is used to decompose the raw data. However, it will be appreciated that other wavelets can be used for the wavelet transformation. Since each stage in a wavelet transformation has one-half the data points of the previous stage, the wavelet transformation can be continued until the stage n low data 89 has around 50 points. Accordingly, the stage n high 88 would contain about 100 data points. Since the preferred wavelet is 24 points long, little data or information can be derived by continuing the wavelet transformation on a data set of around 50 points.

FIG. 28 shows an example of stage 0 high data 95. Since stage 0 high data 95 is generally indicative of the highest frequencies in the mass spectrometry data, stage 0 high data 95 will closely relate to the quantity of high frequency noise in the mass spectrometry data. In FIG. 29, an exponential fitting formula has been applied to the stage 0 high data 95 to generate a stage 0 noise profile 97. In particular, the exponential fitting formula is in the format $A_0 + A_1 EXP (-A_2 m)$. It will be appreciated that other expediential fitting formulas or other types of curve fits may be used.

Referring now to FIG. 30, noise profiles for the other high stages are determined. Since the later data points in each stage will likely be representative of the level of noise in each stage, only the later data points in each stage are used to generate a standard deviation figure that is representative of the noise content in that particular stage. More particularly, in generating the noise profile for each remaining stage, only the last five percent of the data points in each stage are analyzed to determined a standard deviation number. It will be appreciated that other numbers of points, or alternative methods could be used to generate such a standard deviation figure.

15 ⋅

20

25

30

The standard deviation number for each stage is used with the stage 0 noise profile (the exponential curve) 97 to generate a scaled noise profile for each stage. For example, FIG. 30 shows that stage 1 high data 98 has stage 1 high data 103 with the last five percent of the data points represented by area 99. The points in area 99 are evaluated to determine a standard deviation number indicative of the noise content in stage 1 high data 103. The standard deviation number is then used with the stage 0 noise profile 97 to generate a stage 1 noise profile.

In a similar manner, stage 2 high 100 has stage 2 high data 104 with the last five percent of points represented by area 101. The data points in area 101 are then used to calculate a standard deviation number which is then used to scale the stage 0 noise profile 97 to generate a noise profile for stage 2 data. This same process is continued for each of the stage high data as shown by the stage n high 105. For stage n high 105, stage n high data 108 has the last five percent of data points indicated in area 106. The data points in area 106 are used to determine a standard deviation number for stage n. The stage n standard deviation number is then used with the stage 0 noise profile 97 to generate a noise profile for stage n. Accordingly, each of the high data stages has a noise profile.

FIG. 31 shows how the noise profile is applied to the data in each stage. Generally, the noise profile is used to generate a threshold which is applied to the data in each stage. Since the noise profile is already scaled to adjust for the noise content of each stage, calculating a threshold permits further adjustment to tune the quantity of noise removed. Wavelet coefficients below the threshold are ignored while those above the threshold are retained. Accordingly, the remaining data has a substantial portion of the noise content removed.

Due to the characteristics of wavelet transformation, the lower stages, such as stage 0 and 1, will have more noise content than the later stages such as stage 2 or stage-n. Indeed, stage n low data is likely to have little noise at all. Therefore, in a preferred embodiment the noise profiles are applied more aggressively in the lower stages and less aggressively in the later stages. For example, FIG. 31 shows that stage 0 high threshold is determined by multiplying

20

25

30

the stage 0 noise profile by a factor of four. In such a manner, significant numbers of data points in stage 0 high data 95 will be below the threshold and therefore eliminated. Stage 1 high threshold 112 is set at two times the noise profile for the stage 1 high data, and stage 2 high threshold 114 is set equal to the noise profile for stage 2 high. Following this geometric progression, stage n high threshold 116 is therefore determined by scaling the noise profile for each respective stage n high by a factor equal to $(1/2^{n-2})$. It will be appreciated that other factors may be applied to scale the noise profile for each stage. For example, the noise profile may be scaled more or less aggressively to accommodate specific systemic characteristics or sample compositions. As indicated above, stage n low data does not have a noise profile applied as stage n low data 118 is assumed to have little or no noise content. After the scaled noise profiles have been applied to each high data stage, the mass spectrometry data 70 has been denoised and is ready for further processing. A wavelet 15 transformation of the denoised signal results in the sparse data set 120 as shown in FIG. 31.

Referring again to FIG. 25, the mass spectrometry data received in block 40 has been denoised in block 45 and is now passed to block 50 for baseline correction. Before performing baseline correction, the artifacts introduced by the wavelet transformation procedure are preferably removed. Wavelet transformation results vary slightly depending upon which point of the wavelet is used as a starting point. For example, the preferred embodiment uses the 24point Daubechies-24 wavelet. By starting the transformation at the 0 point of the wavelet, a slightly different result will be obtained than if starting at points 1 or 2 of the wavelet. Therefore, the denoised data is transformed using every available possible starting point, with the results averaged to determine a final denoised and shifted signal. For example, FIG. 33 shows that the wavelet coefficient is applied 24 different times and then the results averaged to generate the final data set. It will be appreciated that other techniques may be used to accommodate the slight error introduced due to wavelet shifting.

The formula 125 is generally indicated in FIG. 33. Once the signal has been denoised and shifted, a denoised and shifted signal 130 is generated as

20

25

30

shown in FIG. 58. FIG. 34 shows an example of the wavelet coefficient 1.35 data set from the denoised and shifted signal 130.

FIG. 36 shows that putative peak areas 145, 147, and 149 are located in the denoised and shifted signal 150. The putative peak areas are systematically identified by taking a moving average along the signal 150 and identifying sections of the signal 150 which exceed a threshold related to the moving average. It will be appreciated that other methods can be used to identify putative peak areas in the signal 150.

Putative peak areas 145, 147 and 149 are removed from the signal 150 to create a peak-free signal 155 as shown in FIG. 37. The peak-free signal 155 is further analyzed to identify remaining minimum values 157, and the remaining minimum values 157 are connected to generate the peak-free signal 155.

FIG. 38 shows a process of using the peak-free signal 155 to generate a baseline 170 as shown in FIG. 39. As shown in block 162, a wavelet transformation is performed on the peak-free signal 155. All the stages from the wavelet transformation are eliminated in block 164 except for the n low stage. The n low stage will generally indicate the lowest frequency component of the peak-free signal 155 and therefore will generally indicate the system exponential characteristics. Block 166 shows that a signal is reconstructed from the n low coefficients and the baseline signal 170 is generated in block 168.

FIG. 39 shows a denoised and shifted data signal 172 positioned adjacent a correction baseline 170. The baseline correction 170 is subtracted from the denoised and shifted signal 172 to generate a signal 175 having a baseline correction applied as shown in FIG. 40. Although such a denoised, shifted, and corrected signal is sufficient for most identification purposes, the putative peaks in signal 175 are not identifiable with sufficient accuracy or confidence to call the DNA composition of a biological sample.

Referring again to FIG. 25, the data from the baseline correction 50 is now compressed in block 55, the compression technique used in a preferred embodiment is detailed in FIG. 41. In FIG. 41 the data in the baseline corrected data is presented in an array format 182 with x-axis points 183 having an associated data value 184. The x-axis is indexed by the non-zero wavelet

25

30

coefficients, and the associated value is the value of the wavelet coefficient. In the illustrated data example in table 182, the maximum value 184 is indicated to be 1000. Although a particularly advantageous compression technique for mass spectrometry data is shown, it will be appreciated that other compression techniques can be used. Although not preferred, the data may also be stored without compression.

In compressing the data according to a preferred embodiment, an intermediate format 186 is generated. The intermediate format 186 generally comprises a real number having a whole number portion 188 and a decimal portion 190. The whole number portion is the x-axis point 183 while the decimal portion is the value data 184 divided by the maximum data value. For example, in the data 182 a data value "25" is indicated at x-axis point "100". The intermediate value for this data point would be "100.025".

From the intermediate compressed data 186 the final compressed data 195 is generated. The first point of the intermediate data file becomes the starting point for the compressed data. Thereafter each data point in the compressed data 195 is calculated as follows: the whole number portion (left of the decimal) is replaced by the difference between the current and the last whole number. The remainder (right of the decimal) remains intact. For example, the starting point of the compressed data 195 is shown to be the same as the intermediate data point which is "100.025". The comparison between the first intermediate data point "100.025" and the second intermediate data point "150.220" is "50.220". Therefore, "50.220" becomes the second point of the compressed data 195. In a similar manner, the second intermediate point is "150.220" and the third intermediate data point is "500.0001". Therefore, the third compressed data becomes "350.000". The calculation for determining compressed data points is continued until the entire array of data points is converted to a single array of real numbers.

FIG. 42 generally describes the method of compressing mass spectrometry data, showing that the data file in block 201 is presented as an array of coefficients in block 202. The data starting point and maximum is a determined as shown in block 203, and the intermediate real numbers are

25

calculated in block 204 as described above. With the intermediate data points generated, the compressed data is generated in block 205. The described compression method is highly advantageous and efficient for compressing data sets such as a processed data set from a mass spectrometry instrument. The method is particularly useful for data, such as mass spectrometry data, that uses large numbers and has been processed to have occasional lengthy gaps in x-axis data. Accordingly, an x-y data array for processed mass spectrometry data may be stored with an effective compression rate of 10x or more. Although the compression technique is applied to mass spectrometry data, it will be appreciated that the method may also advantageously be applied to other data sets.

Referring again to FIG. 25, peak heights are now determined in block 60. The first step in determining peak height is illustrated in FIG. 43 where the signal 210 is shifted left or right to correspond with the position of expected peaks. As the set of possible compositions in the biological sample is known before the mass spectrometry data is generated, the possible positioning of expected peaks is already known. These possible peaks are referred to as expected peaks, such as expected peaks 212, 214, and 216. Due to calibration or other errors in the test instrument data, the entire signal may be shifted left or right from its actual position, therefore, putative peaks located in the signal, such as putative peaks 218, 222, and 224 may be compared to the expected peaks 212, 214, and 216, respectively. The entire signal is then shifted such that the putative peaks align

Once the putative peaks have been shifted to match expected peaks, the strongest putative peak is identified in FIG. 44. In a preferred embodiment, the strongest peak is calculated as a combination of analyzing the overall peak height and area beneath the peak. For example, a moderately high but wide peak would be stronger than a very high peak that is extremely narrow. With the strongest putative peak identified, such as putative peak 225, a Gaussian 228 curve is fit to the peak 225. Once the Gaussian is fit, the width (W) of the Gaussian is determined and will be used as the peak width for future calculations.

more closely with the expected peaks.

20

25

30

As generally addressed above, the denoised, shifted, and baseline-corrected signal is not sufficiently processed for confidently calling the DNA composition of the biological sample. For example, although the baseline has generally been removed, there are still residual baseline effects present. These residual baseline effects are therefore removed to increase the accuracy and confidence in making identifications.

To remove the residual baseline effects, FIG. 45 shows that the putative peaks 218, 222, and 224 are removed from the baseline corrected signal. The peaks are removed by identifying a center line 230, 232, and 234 of the putative peaks 218, 222, and 224, respectively and removing an area to the left and to the right of the identified center line. For each putative peak, an area equal to twice the width (W) of the Gaussian is removed from the left of the center line, while an area equivalent to 50 daltons is removed from the right of the center line. It has been found that the area representing 50 daltons is adequate to sufficiently remove the effect of salt adducts which may be associated with an actual peak. Such adducts appear to the right of an actual peak and are a natural effect from the chemistry involved in acquiring a mass spectrum. Although a 50 Dalton buffer has been selected, it will be appreciated that other ranges or methods can be used to reduce or eliminate adduct effects.

The peaks are removed and remaining minima 247 located as shown in FIG. 46 with the minima 247 connected to create signal 245. A quartic polynomial is applied to signal 245 to generate a residual baseline 250 as shown in FIG. 47. The residual baseline 250 is subtracted from the signal 225 to generate the final signal 255 as indicated in FIG. 48. Although the residual baseline is the result of a quartic fit to signal 245, it will be appreciated that other techniques can be used to smooth or fit the residual baseline.

To determine peak height, as shown in FIG. 49, a Gaussian such as Gaussian 266, 268, and 270 is fit to each of the peaks, such as peaks 260, 262, and 264, respectively. Accordingly, the height of the Gaussian is determined as height 272, 274, and 276. Once the height of each Gaussian peak is determined, then the method of identifying a biological compound 35, can move into the genotyping phase 65 as shown in FIG. 25.

20

25

30

An indication of the confidence that each putative peak is an actual peak can be discerned by calculating a signal-to-noise ratio for each putative peak. Accordingly, putative peaks with a strong signal-to-noise ratio are generally more likely to be an actual peak than a putative peak with a lower signal-to-noise ratio. As described above and shown in FIG. 50, the height of each peak, such as height 272, 274, and 276, is determined for each peak, with the height being an indicator of signal strength for each peak. The noise profile, such as noise profile 97, is extrapolated into noise profile 280 across the identified peaks. At the center line of each of the peaks, a noise value is determined, such as noise value 282, 283, and 284. With a signal values and a noise values generated, signal-to-noise ratios can be calculated for each peak. For example, the signal-to-noise ratio for the first peak in FIG. 50 would be calculated as signal value 272 divided by noise value 282, and in a similar manner the signal-to-noise ratio of the middle peak in FIG. 50 would be determined as signal 274 divided by noise value 283.

Although the signal-to-noise ratio is generally a useful indicator of the presence of an actual peak, further processing has been found to increase the confidence by which a sample can be identified. For example, the signal-to-noise ratio for each peak in the preferred embodiment is preferably adjusted by the goodness of fit between a Gaussian and each putative peak. It is a characteristic of a mass spectrometer that sample material is detected in a manner that generally complies with a normal distribution. Accordingly, greater confidence will be associated with a putative signal having a Gaussian shape than a signal that has a less normal distribution. The error resulting from having a non-Gaussian shape can be referred to as a "residual error".

Referring to FIG. 51, a residual error is calculated by taking a root mean square calculation between the Gaussian 293 and the putative peak 290 in the data signal. The calculation is performed on data within one width on either side of a center line of the Gaussian. The residual error is calculated as:

where G is the Gaussian signal value, R is the putative peak value, and N is the number of points from -W to +W. The calculated residual error is used to generate an adjusted signal-to-noise ratio, as described below.

25

An adjusted signal noise ratio is calculated for each putative peak using the formula (S/N) * EXP^(-,1,R), where S/N is the signal-to-noise ratio, and R is the residual error determined above. Although the preferred embodiment calculates an adjusted signal-to-noise ratio using a residual error for each peak, it will be appreciated that other techniques can be used to account for the goodness of fit between the Gaussian and the actual signal.

Referring now to FIG. 52, a probability is determined that a putative peak is an actual peak. In making the determination of peak probability, a probability profile 300 is generated where the adjusted signal-to-noise ratio is the x-axis and the probability is the y-axis. Probability is necessarily in the range between a 0% probability and a 100% probability, which is indicated as 1. Generally, the higher the adjusted signal-to-noise ratio, the greater the confidence that a putative peak is an actual peak.

At some target value for the adjusted signal-to-noise, it has been found that the probability is 100% that the putative peak is an actual peak and can confidently be used to identify the DNA composition of a biological sample. However, the target value of adjusted signal-to-noise ratio where the probability is assumed to be 100% is a variable parameter which is to be set according to application specific criteria. For example, the target signal-to-noise ratio will be adjusted depending upon trial experience, sample characteristics, and the acceptable error tolerance in the overall system. More specifically, for situations requiring a conservative approach where error cannot be tolerated, the target adjusted signal-to-noise ratio can be set to, for example, 10 and higher. Accordingly, 100% probability will not be assigned to a peak unless the adjusted signal-to-noise ratio is 10 or over.

In other situations, a more aggressive approach may be taken as sample data is more pronounced or the risk of error may be reduced. In such a situation, the system may be set to assume a 100% probability with a 5 or greater target signal-to-noise ratio. Of course, an intermediate signal-to-noise ratio target figure can be selected, such as 7, when a moderate risk of error can be assumed. Once the target adjusted signal-to-noise ratio is set for the method,

20

25

30

then for any adjusted signal-to-noise ratio a probability can be determined that a putative peak is an actual peak.

Due to the chemistry involved in performing an identification test, especially a mass spectrometry test of a sample prepared by DNA amplifications, the allelic ratio between the signal strength of the highest peak and the signal strength of the second (or third and so on) highest peak should fall within an expected ratio. If the allelic ratio falls outside of normal guidelines, the preferred embodiment imposes an allelic ratio penalty to the probability. For example, FIG. 53 shows an allelic penalty 315 which has an x-axis 317 that is the ratio between the signal strength of the second highest peak divided by signal strength of the highest peak. The y-axis 319 assigns a penalty between 0 and 1 depending on the determined allelic ratio. In the preferred embodiment, it is assumed that allelic ratios over 30% are within the expected range and therefore no penalty is applied. Between a ratio of 10% and 30%, the penalty is linearly increased until at allelic ratios below 10% it is assumed the second-highest peak is not real. For allelic ratios between 10% and 30%, the allelic penalty chart 315 is used to determine a penalty 319, which is multiplied by the peak probability determined in FIG. 52 to determine a final peak probability. Although the preferred embodiment incorporates an allelic ratio penalty to account for a possible chemistry error, it will be appreciated that other techniques may be used. Similar treatment will be applied to the other peaks.

With the peak probability of each peak determined, the statistical probability for various composition components may be determined. As an example, in order to determine the probability of each of three possible combinations of two peaks, -- peak G, peak C and combinations GG, CC and GC. FIG. 54 shows an example where a most probable peak 325 is determined to have a final peak probability of 90%. Peak 325 is positioned such that it represents a G component in the biological sample. Accordingly, it can be maintained that there is a 90% probability that G exists in the biological sample. Also in the example shown in FIG. 54, the second highest probability is peak 330 which has a peak probability of 20%. Peak 330 is at a position associated

20

25

30

with a C composition. Accordingly, it can be maintained that there is a 20% probability that C exists in the biological sample.

With the probability of G existing (90%) and the probability of C existing (20%) as a starting point, the probability of combinations of G and C existing can be calculated. For example, FIG. 54 indicates that the probability of GG existing 329 is calculated as 72%. This is calculated as the probability of GG is equal to the probability of G existing (90%) multiplied by the probability of C not existing (100% -20%). So if the probability of G existing is 90% and the probability of C not existing is 80%, the probability of GG is 72%.

In a similar manner, the probability of CC existing is equivalent to the probability of C existing (20%) multiplied by the probability of G not existing (100% - 90%). As shown in FIG. 54, the probability of C existing is 20% while the probability of G not existing is 10%, so therefore the probability of CC is only 2%. Finally, the probability of GC existing is equal to the probability of G existing (90%) multiplied by the probability of C existing (20%). So if the probability of G existing is 90% and the probability of C existing is 20%, the probability of GC existing is 18%. In summary form, then, the probability of the composition of the biological sample is:

probability of GG: 72%;

probability of GC: 18%; and

probability of CC: 2%.

Once the probabilities of each of the possible combinations has been determined, FIG. 55 is used to decide whether or not sufficient confidence exists to call the genotype. FIG. 55 shows a call chart 335 which has an x-axis 337 which is the ratio of the highest combination probability to the second highest combination probability. The y-axis 339 simply indicates whether the ratio is sufficiently high to justify calling the genotype. The value of the ratio may be indicated by M 340. The value of M is set depending upon trial data, sample composition, and the ability to accept error. For example, the value M may be set relatively high, such as to a value 4 so that the highest probability must be at least four times greater than the second highest probability before confidence is established to call a genotype. However, if a certain level of error may be

20

25

acceptable, the value of M may be set to a more aggressive value, such as to 3, so that the ratio between the highest and second highest probabilities needs to be only a ratio of 3 or higher. Of course, moderate value may be selected for M when a moderate risk can be accepted. Using the example of FIG. 54, where the probability of GG was 72% and the probability of GC was 18%, the ratio between 72% and 18% is 4.0, therefore, whether M is set to 3, 3.5, or 4, the system would call the genotype as GG. Although the preferred embodiment uses a ratio between the two highest peak probabilities to determine if a genotype confidently can be called, it will be appreciated that other methods may be substituted. It will also be appreciated that the above techniques may be used for calculating probabilities and choosing genotypes (or more general DNA patterns) containing of combinations of more than two peaks.

Referring now to FIG. 56, a flow chart is shown generally defining the process of statistically calling genotype described above. In FIG. 56 block 402 shows that the height of each peak is determined and that in block 404 a noise profile is extrapolated for each peak. The signal is determined from the height of each peak in block 406 and the noise for each peak is determined using the noise profile in block 408. In block 410, the signal-to-noise ratio is calculated for each peak. To account for a non-Gaussian peak shape, a residual error is determined in block 412 and an adjusted signal-to-noise ratio is calculated in block 414. Block 416 shows that a probability profile is developed, with the probability of each peak existing found in block 418. An allelic penalty may be applied in block 420, with the allelic penalty applied to the adjusted peak probability in block 422. The probability of each combination of components is calculated in block 424 with the ratio between the two highest probabilities being determined in block 426. If the ratio of probabilities exceeds a threshold value then the genotype is called in block 428.

In another embodiment of the invention, the computing device 20 (Fig. 24) supports "standardless" genotyping by identifying data peaks that contain putative SNPs. Standardless genotyping is used, for example, where insufficient information is known about the samples to determine a distribution of expected

20

25

peak locations, against which an allelic penalty as described above can be reliably calculated. This permits the computing device to be used for identification of peaks that contain putative SNPs from data generated by any assay that fragments a targeted DNA molecule. For such standardless genotyping, peaks that are associated with an area under the data curve that deviates significantly from the typical area of other peaks in the data spectrum are identified and their corresponding mass (location along the x-axis) is determined.

More particularly, peaks that deviate significantly from the average area of other peaks in the data are identified, and the expected allelic ratio between data peaks is defined in terms of the ratio of the area under the data peaks. Theoretically, where each genetic loci has the same molar concentration of analyte, the area under each corresponding peak should be the same, thus producing a 1.0 ratio of the peak area between any two peaks. In accordance with the invention, peaks having a smaller ratio relative to the other peaks in the data will not be recognized as peaks. More particularly, peaks having an area ratio smaller than 30% relative to a nominal value for peak area will be assigned an allelic penalty. The mass of the remaining peaks (their location along the x-axis of the data) will be determined based on oligonucleotide standards.

Fig. 57 shows a flow diagram representation of the processing by the computing device 20 (Fig. 24) when performing standardless genotyping. In the first operation, represented by the flow diagram box numbered 502, the computing device receives data from the mass spectrometer. Next, the height of each putative peak in the data sample is determined, as indicated by the block 504. After the height of each peak in the mass spectrometer data is determined, a de-noise process 505 is performed, beginning with an extrapolation of the noise profile (block 506), followed by finding the noise of each peak (block 508) and calculating the signal to noise ratio for each data sample (block 510). Each of these operations may be performed in accordance with the description above for denoise operations 45 of Fig. 25. Other suitable denoise operations will occur to those skilled in the art.

20

25

The next operation is to find the residual error associated with each data point. This is represented by the block 512 in Figure 57. The next step, block 514, involves calculating an adjusted signal to noise ratio for each identified peak. A probability profile is developed next (block 516), followed by a determination of the peak probabilities at block 518. In the preferred embodiment, the denoise operations of Fig. 57, comprising block 502 to block 518, comprise the corresponding operations described above in conjunction with Fig. 56 for block 402 through block 418, respectively.

The next action for the standardless genotype processing is to determine an allelic penalty for each peak, indicated by the block 524. As noted above, the standardless genotype processing of Fig. 57 determines an allelic penalty by comparing area under the peaks. Therefore, rather than compare signal strength ratios to determine an allelic penalty, such as described above for Fig. 53, the standardless processing determines the area under each of the identified peaks and compares the ratio of those areas. Determining the area under each peak may be computed using conventional numerical analysis techniques for calculating the area under a curve for experimental data.

Thus, the allelic penalty is assigned in accordance with Fig. 58, which shows that no penalty is assigned to peaks having a peak area relative to an expected average area value that is greater than 0.30 (30%). The allelic penalty is applied to the peak probability value, which may be determined according to the process such as described in Fig. 52. It should be apparent from Fig. 58 that the allelic penalty imposed for peaks below a ratio of 30% is that such peaks will be removed from further measurement and processing. Other penalty schemes, however, may be imposed in accordance with knowledge about the data being processed, as determined by those skilled in the art.

After the allelic penalty has been determined and applied, the standardless genotype processing compares the location of the remaining putative peaks to oligonucleotide standards to determine corresponding masses in the processing for block 524. For standardless genotype data, the processing of the block 524 is performed to determine mass and genotype, rather than performing the operations corresponding to block 424, 426, and 428 of Fig. 33.

10

15

20

25

30

Techniques for performing such comparisons and determining mass will be known to those skilled in the art.

In another embodiment, the computing device 20 (Fig. 24) permits the detection and determination of the mass (location along the x-axis of the data) of the sense and antisense strand of fragments generated in the assay. If desired, the computing device may also detect and determine the quantity (area under each peak) of the respective sense and antisense strands, using a similar technique to that described above for standardless genotype processing. The data generated for each type of strand may then be combined to achieve a data redundancy and to thereby increase the confidence level of the determined genotype. This technique obviates primer peaks that are often observed in data from other diagnostic methods, thereby permitting a higher level of multiplexing. In addition, when quantitation is used in pooling experiments, the ratio of the measured peak areas is more reliably calculated than the peak identifying technique, due to data redundancy.

Fig. 23 is a flow diagram that illustrates the processing implemented by the computing device 20 to perform sense and antisense processing. In the first operation, represented by the flow diagram box numbered 602, the computing device receives data from the mass spectrometer. This data will include data for the sense strand and antisense strand of assay fragments. Next, the height of each putative peak in the data sample is determined, as indicated by the block 604. After the height of each peak in the mass spectrometer data is determined, a de-noise process 605 is performed, beginning with an operation that extrapolates the noise profile (block 606), followed by finding the noise of each peak (block 608) and calculating the signal to noise ratio for each data sample (block 610). Each of these operations may be performed in accordance with the description above for the denoise operations 45 of Fig. 25. Other suitable denoise operations will occur to those skilled in the art. The next operation is to find the residual error associated with each data point. This is represented by the block 612 in Figure 36.

After the residual error for the data of the sense strand and antisense strand has been performed, processing to identify the genotypes will be

performed for the sense strand and also for the antisense strand. Therefore, Fig. 23 shows that processing includes sense strand processing (block 630) and antisense strand processing (block 640). Each block 630, 640 includes processing that corresponds to adjusting the signal to noise ratio, developing a probability profile, determining an allelic penalty, adjusting the peak probability by the allelic penalty, calculating genotype probabilities, and testing genotype probability ratios, such as described above in conjunction with blocks 414 through 426 of Fig. 56. The processing of each block 630, 640 may, if desired, include standardless processing operations such as described above in conjunction with Fig. 57. The standardless processing may be included in place of or in addition to the processing operations of Fig. 56.

After the genotype probability processing is completed, the data from the sense strand and antisense strand processing is combined and compared to expected database values to obtain the benefits of data redundancy as between the sense strand and antisense strand. Those skilled in the art will understand techniques to take advantage of known data redundancies between a sense strand and antisense strand of assay fragments. This processing is represented by the block 650. After the data from the two strands is combined for processing, the genotype processing is performed (block 660) and the genotype is identified.

Since modifications will be apparent to those of skill in this art, it is intended that this invention be limited only by the scope of the appended claims.

WHAT IS CLAIMED IS:

A subcollection of samples from a target population, comprising:
 a plurality of samples, wherein the samples are selected from the group
 consisting of blood, tissue, body fluid, cell, seed, microbe, pathogen and
 reproductive tissue samples; and

a symbology on the containers containing the samples, wherein the symbology is representative of the source and/or history of each sample, wherein:

the target population is a healthy population that has not been selected 10 for any disease state;

the collection comprises samples from the healthy population; and the subcollection is obtained by sorting the collection according to specified parameters.

- The subcollection of claim 1, wherein the parameters are selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of a disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental factors.
 - 3. The subcollection of claim 1, wherein the symbology is a bar code.
 - 4. A method of producing a database, comprising:

identifying healthy members of a population;

obtaining data comprising identifying information and obtaining historical information and data relating to the identified members of the population and their immediate family;

entering the data into a database for each member of the population and associating the member and the data with an indexer.

The method of claim 4, further comprising:
 obtaining a body tissue or body fluid sample;

15

20

25

30

analyzing the body tissue or body fluid in the sample; and entering the results of the analysis for each member into the database and associating each result with the indexer representative of each member.

- 6. A database produced by the method of claim 4.
- 7. A database produced by the method of claim 5.
- 8. A database, comprising:

datapoints representative of a plurality of healthy organisms from whom biological samples are obtained,

wherein each datapoint is associated with data representative of the organism type and other identifying information.

- 9. The database of claim 8, wherein the datapoints are answers to questions regarding one or more of a parameters selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of a disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental factors.
- 10. The database of claim 9, wherein the organisms are mammals and the samples are body fluids or tissues.
- 11. The database of claim 9, wherein the samples are selected from blood, blood fractions, cells and subcellular organelles.
 - 12 The database of claim 8, further comprising, phenotypic data from an organism.
- 13. The database of claim 12, wherein the data includes one of physical characteristics, background data, medical data, and historical data.
- 14. The database of claim 8, further comprising, genotypic data from nucleic acid obtained from an organism.

20

25

- 15. The database of claim 14, wherein genotypic data includes, genetic markers, non-coding regions, microsatellites, RFLPs, VNTRs, historical data of the organism, medical history, and phenotypic information.
 - 16. The database of claim 8 that is a relational database.
- 17. The database of claim 16, wherein the data are related to an indexer datapoint representative of each organism from whom data is obtained.
- 18. A method of identifying polymorphisms that are candidate genetic markers, comprising:

identifying a polymorphism; and

identifying any pathway or gene linked to the locus of the polymorphism, wherein

the polymorphisms are identified in samples associated with a target population that comprises healthy subjects.

- 19. The method of claim 18, wherein the polymorphism is identified by detecting the presence of target nucleic acids in a sample by a method, comprising the steps of:
 - a) hybridizing a first oligonucleotide to the target nucleic acid;
 - b) hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;
 - c) ligating the hybridized oligonucleotides; and
 - c) detecting hybridized first oligonucleotide by mass spectrometry as an indication of the presence of the target nucleic acid.
 - 20. The method of claim 18, wherein the polymorphism is identified by detecting target nucleic acids in a sample by a method, comprising the steps of:
 - a) hybridizing a first oligonucleotide to the target nucleic acid and hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;
 - b) contacting the hybridized first and second oligonucleotides with a cleavage enzyme to form a cleavage product; and
- 30 c) detecting the cleavage product by mass spectrometry as an indication of the presence of the target nucleic acid.

20

25

- 21. The method of claim 20 wherein the samples are from subjects in a healthy database.
- 22. The method of claim 18, wherein the polymorphism is identified by identifying target nucleic acids in a sample by primer oligo base extension (probe).
- 23. The method of 22, wherein primer oligo base extension, comprises:
 - a) obtaining a nucleic acid molecule that contains a target nucleotide;
- b) optionally immobilizing the nucleic acid molecule onto a solid support,
 to produce an immobilized nucleic acid molecule;
 - c) hybridizing the nucleic acid molecule with a primer oligonucleotide that is complementary to the nucleic acid molecule at a site adjacent to the target nucleotide;
 - d) contacting the product of step c) with a composition comprising a dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate that is complementary to the target nucleotide is extended onto the primer; and
 - e) detecting the extended primer, thereby identifying the target nucleotide.
 - 24. The method of claim 23, wherein detection of the extended primer is effected by mass spectrometry, comprising:

ionizing and volatizing the product of step d); and detecting the extended primer by mass spectrometry, thereby identifying the target nucleotide.

25. The method of claim 24, wherein; samples are presented to the mass spectrometer as arrays on chips; and each sample occupies a volume that is about the size of the laser spot projected by the laser in a mass spectrometer used in matrix-assisted laser desorption/ionization (MALDI) spectrometry.

10

15

20

25

30

26. A combination, comprising:

a database containing parameters associated with a datapoint representative of a subject from whom samples are obtained, wherein the subjects are healthy; and

an indexed collection of the samples, wherein the index identifies the subject from whom the sample was obtained.

- 27 The combination of claim 26, wherein the parameter is selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of disease in immediate family (parent, siblings, children), use of prescription drugs and/or reason therefor, length and/or number of hospital stays and ecposure to environmental factors.
- 28. The combination of claim 26, wherein the database further contains genotypic data for each subject.
 - 29. The combination of claim 26, wherein the samples are blood.
 - 30 A data storage medium, comprising the database of claim 8.
 - 31. A computer system, comprising the database of claim 8.
- 32. A system for high throughput processing of biological samples, comprising:
 - a process line comprising a plurality of processing stations, each of which performs a procedure on a biological sample contained in a reaction vessel;
 - a robotic system that transports the reaction vessel from processing station to processing station;
 - a data analysis system that receives test results of the process line and automatically processes the test results to make a determination regarding the biological sample in the reaction vessel;
 - a control system that determines when the test at each processing station is complete and, in response, moves the reaction vessel to

15

20

25

the next test station, and continuously processes reaction vessels one after another until the control system receives a stop instruction; and

a database of claim 8, wherein the samples tested by the automated process line comprise samples from subjects in the database.

- 33. The system of claim 32, wherein one of the processing stations comprises a mass spectrometer.
- 34. The system of claim 32, wherein the data analysis system processes the test results by receiving test data from the mass spectrometer such that the test data for a biological sample contains one or more signals, whereupon the data analysis system determines the area under the curve of each signal and normalizes the results thereof and obtains a substantially quantitative result representative of the relative amounts of components in the tested sample.
- 35. A method for high throughput processing of biological samples, the method comprising:
 - transporting a reaction vessel along a system of claim 32, comprising a process line having a plurality of processing stations, each of which performs a procedure on one or more biological samples contained in the reaction vessel;
 - determining when the test procedure at each processing station is complete and, in response, moving the reaction vessel to the next processing station;
 - receiving test results of the process line and automatically processing the test results to make a data analysis determination regarding the biological samples in the reaction vessel; and
 - processing reaction vessels continuously one after another until receiving a stop instruction, wherein the samples tested by the automated process line comprise samples from subjects in the database.
- 30 36. The method of 35, wherein one of the processing stations comprises a mass spectrometer.

20

30

- 37. The method of claim 36, wherein the samples are analyzed by a method comprising primer oligo base extension (probe).
 - 38. The method of claim 37, further comprising:

processing the test results by receiving test data from the mass spectrometer such that the test data for a biological sample contains one or more signals or numerical values representative of signals, whereupon the data analysis system determines the area under the curve of each signal and normalizes the results thereof and obtains a substantially quantitative result representative of the relative amounts of components in the tested sample.

- 39. The method of claim 37, wherein primer oligo base extension, comprises:
 - a) obtaining a nucleic acid molecule that contains a target nucleotide;
- b) optionally immobilizing the nucleic acid molecule onto a solid support, to produce an immobilized nucleic acid molecule;
- 15 c) hybridizing the nucleic acid molecule with a primer oligonucleotide that is complementary to the nucleic acid molecule at a site adjacent to the target nucleotide;
 - d) contacting the product of step c) with composition comprising a dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate that is complementary to the target nucleotide is extended onto the primer; and
 - e) detecting the primer, thereby identifying the target nucleotide.
 - 40. The method of 39, wherein detection of the extended primer is effected by mass spectrometry, comprising:

25 ionizing and volatizing the product of step d); and detecting the extended primer by mass spectrometry, thereby identifying the target nucleotide.

- 41. The method of claim 36, wherein the target nucleic acids in the sample are detected and/or identified by a method, comprising the steps of:
 - a) hybridizing a first oligonucleotide to the target nucleic acid;
- b) hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;

20

25

- c) ligating then hybridized oligonucleotides; and
- c) detecting hybridized first oligonucleotide by mass spectrometry as an indication of the presence of the target nucleic acid.
- 42. The method of claim 36, wherein the target nucleic acids in the sample are detected and/or identified by a method, comprising the steps of:
- a) hybridizing a first oligonucleotide to the target nucleic acid and hybridizing a second oligonucleotide to an adjacent region of the target nucleic acid;
- b) contacting the hybridized first and second oligonucleotides with a
 O cleavage enzyme to form a cleavage product; and
 - c) detecting the cleavage product by mass spectrometry as an indication of the presence of the target nucleic acid.
 - 43. A method of producing a database stored in a computer memory, comprising:

identifying healthy members of a population;

obtaining identifying and historical information and data relating to the identified members of the population;

entering the member-related data into the computer memory database for each identified member of the population and associating the member and the data with an indexer.

- 44. The method of claim 43, further comprising:
 obtaining a body tissue or body fluid sample of an identified member;
 analyzing the body tissue or body fluid in the sample; and
 entering the results of the analysis for each member into the computer
 memory database and associating each result with the indexer representative of
 each member.
 - 45. A database produced by the method of claim 43.
 - 46. A database produced by the method of claim 44.
 - 47. The database of claim 8, wherein:
- the organims are selected from among animals, bacteria, fungi, protozoans and parasites and

15

20

each datapoint is associated with parameters representative of the organism type and identifying information.

- 48. The database of claim 43, further comprising, phenotypic data regarding each subject.
- 49. The database of claim 47 that is a relational database and the parameters are the answers to the questions in the questionnaire.
 - 50. The database of claim 8, further comprising,

genotypic data of nucleic acid of the subject, wherein genotypic data includes, but is not limited to, genetic markers, non-coding regions,

- microsatellites, restriction fragment length polymorphisms (RFLPs), variable number tandem repeats (VNTRs), historical day of the organism, the medical history of the subject, phenotypic information, and other information.
 - 51. A database, comprising data records stored in computer memory, wherein the data records contain information that identifies healthy members of a population, and also contain identifying and historical information and data relating to the identified members.
 - 52. The database of claim 51, further comprising an index value for each identified member that associates each member of the population with the identifying and historical information and data.
 - 53. A computer system, comprising the database of claim 51.
 - 54. An automated process line, comprising the database of claim 51.
 - 55. A method for determining a polymorphism that correlates with age, ethnicity or gender, comprising:

identifying a polymorphism; and

- determining the frequency of the polymorphism with increasing age, with ethnicity or with gender in a healthy population.
 - 56. A method for determining whether a polymorphism correlates with suceptibility to morbidity, early mortality, or morbidity and early mortality, comprising;
- identifying a polymorphism; and determining the frequency of the polymorphism with increasing age in a healthy population.

15

20

25

57. A high throughput method of determining frequencies of genetic variations, comprising:

selecting a healthy target population and a genetic variation to be assessed;

pooling a plurality of samples of biopolymers obtained from members of the population,

determining or detecting the biopolymer that comprises the variation by mass spectrometry;

obtaining a mass spectrum or a digital representation thereof; and determining the frequency of the variation in the population.

58. The method of claim 57, wherein:

the variation is selected from the group consisting of an allelic variation, a post-translational modification, a nucleic modification, a label, a mass modification of a nucleic acid and methylation; and/or

the biopolymer is a nucleic acid, a protein, a polysaccharide, a lipid, a small organic metabolite or intermediate, wherein the concentration of biopolymer of interest is the same in each of the samples; and/or

the frequency is determined by assessing the method comprising determining the area under the peak in the mass spectrum or digital repesentation thereof corresponding to the mass of the biopolymer comprising the genomic variation.

- 59. The method of claim 58, wherein the method for determining the frequency is effected by determining the ratio of the signal or the digital representation thereof to the total area of the entire mass spectrum, which is corrected for background.
- 60. A method for discovery of a polymorphism in a population, comprising:

sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

isolating a biopolymer from each identified sample; optionally pooling each isolated biopolymer; optionally amplifying the amount of biopolymer;

25

cleaving the pooled biopolymers to produce fragments thereof; obtaining a mass spectrum of the resulting fragments and comparing the mass spectrum with a control mass spectrum to identify differences between the spectra and thereby identifing any polymorphisms; wherein:

the control mass spectrum is obtained from unsorted samples in the collection or samples sorted according to a different parameter.

- 61. The method of claim 60, wherein cleaving is effected by contacting the biopolymer with an enzyme.
- The method of claim 61, wherein the enzyme is selected from the group consisting of nucleotide glycosylase, a nickase and a type IIS restriction enzyme.
 - The method of claim 60, wherein the biopolymer is a nucleic acid 63. or a protein.
- The method of claim 60, wherein the the mass spectrometric 64. format is selected from among Matrix-Assisted Laser Desorption/Ionization, 15 Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof.
 - A method for discovery of a polymorphism in a population, comprising:
- obtaining samples of body tissue or fluid from a plurality of organisms; 20 isolating a biopolymer from each sample; pooling each isolated biopolymer; optionally amplifying the amount of biopolymer; cleaving the pooled biopolymers to produce fragments thereof; obtaining a mass spectrum of the resulting fragments;
 - comparing the frequency of each fragment to identify fragments present in amounts lower than the average frequency, thereby identifying any polymorphisms.
- 66. The method of claim 65, wherein cleaving is effected by contacting the biopolymer with an enzyme. 30

25

30

- 67. The method of claim 66, wherein the enzyme is selected from the group consisting of nucleotide glycosylase, a nickase and a type IIS restriction enzyme.
- 68. The method of claim 65, wherein the biopolymer is a nucleic acid or a protein.
 - 69. The method of claim 65, wherein the the mass spectrometric format is selected from among Matrix-Assisted Laser Desorption/Ionization, Time-of-Flight (MALDI-TOF), Electrospray (ES), IR-MALDI, Ion Cyclotron Resonance (ICR), Fourier Transform and combinations thereof.
- 10 70. The method of claim 65, wherein the samples are obtained from healthy subjects.
 - 71. A method of correlating a polymorphism with a parameter, comprising:

sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

isolating a biopolymer from each identified sample;
pooling each isolated biopolymer;
optionally amplifying the amount of biopolymer;
determining the frequency of the polymorphism in the pooled

biopolymers, wherein:

an alteration of the frequency of the polymorphism compared to a control, indicates a correlation of the polymorphism with the selected parameter; and

the control is the frequency of the polymorphism in pooled biopolymers obtained from samples identified from an unsorted database or from a database sorting according to a different parameter.

72. The method claim 71, wherein the parameter is selected from the group consisting of ethnicity, age, gender, height, weight, alcohol intake, number of pregnancies, number of live births, vegetarians, type of physical activity, state of residence and/or length of residence in a particular state, educational level, age of parent at death, cause of parent death, former or current smoker, length of time as a smoker, frequency of smoking, occurrence of a disease in immediate family (parent, siblings, children), use of prescription

PCT/US00/28413

10

15

25

30

drugs and/or reason therefor, length and/or number of hospital stays and exposure to environmental factors.

- 73. The method claim 72, wherein the parameter is occurrence of disease or a particular disease in an immediate family member, thereby correlating the polymorphism with the disease.
- 74. The method of claim 71, wherein the pooled biopolymers are pooled nucleic acid molecules.
- 75. The method of claim 74, wherein the polymorphism is detected by primer oligo base extension (PROBE).
- 76. The method of 75, wherein primer oligo base extension, comprises:
- a) optionally immobilizing the nucleic acid molecules onto a solid support,
 to produce immobilized nucleic acid molecules;
- b) hybridizing the nucleic acid molecules with a primer oligonucleotide that is complementary to the nucleic acid molecule at a site adjacent to the polymorphism;
- c) contacting the product of step c) with composition comprising a dideoxynucleoside triphosphate or a 3'-deoxynucleoside triphosphates and a polymerase, so that only a dideoxynucleoside or 3'-deoxynucleoside triphosphate that is complementary to the polymorphism is extended onto the primer; and
- d) detecting the extended primer, thereby detecting the polymorphism in nucleic acid molecules in the pooled nucleic acids.
- 77. The method of claim 76, wherein detecting is effected by mass spectrometry.
- 78. The method of claim 71, wherein the frequency is percentage of nucleic acid molecules in the pooled nucleic acids that contain the polymorphism.
- 79. The method of claim 78, wherein the ratio is determined by obtaining mass spectra of the pooled nucleic acids.
- 80. The method of claim 72, wherein the parameter is age, thereby correlating the polymorphism with suceptibility to morbidity, early mortality or morbidity and early mortality.

- 81. A method for haplotyping polymorphisms in a nucleic acid, comprising:
- (a) sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;
 - (b) isolating nucleic acid from each identified sample;
 - (c) optionally pooling each isolated nucleic acid;
 - (d) amplifying the amount of nucleic acid;
- (e) forming single-stranded nucleic acid and splitting each singlestrand into a separate reaction vessel;
- 10 (f) contacting each single-stranded nucleic acid with an adaptor nucleic acid to form an adaptor complex;
 - (g) contacting the adaptor complex with a nuclease and a ligase;
 - (h) contacting the products of step (g) with a mixture that is capable of amplifying a ligated adaptor to produce an extended product;
- (i) obtaining a mass spectrum of each nucleic acid resulting from step
 (h) and detecting a polymorphism by identifying a signal corresponding to the extended product;
 - (j) repeating steps (f) through (i) utilizing an adaptor nucleic acid able to hybridize with another adapter nucleic acid that hybridizes to a different sequence on the same strand; whereby

the polymorphisms are haplotyped by detecting more than one extended product.

- 82. The method of claim 1, wherein the nuclease is Fen-1.
- 83. A method for haplotyping polymorphisms in a population,
- 25 comprising:

20

30

sorting the database of claim 8 according to a selected parameter to identify samples that match the selected parameter;

isolating a nucleic acid from each identified sample; pooling each isolated nucleic acid;

optionally amplifying the amount of nucleic acid;

contacting the nucleic acid with at least one enzyme to produce fragments thereof;

obtaining a mass spectrum of the resulting fragments; whereby:

the polymorphisms are detected by detecting signals corresponding to the
polymorphisms; and

the polymorphisms are haplotyped by determining from the mass spectrum that the polymorphisms are located on the same strand of the nucleic acid.

- 84. The method of claim 83, wherein the enzyme is a nickase.
- 85. The method of claim 84, wherein the nickase is selected from the group consisting of NY2A and NYS1.
- 10 86. A method for detecting methylated nucleotides within a nucleic acid sample, comprising:

splitting a nucleic acid sample into separate reaction vessels; contacting nucleic acid in one reaction vessel with bisulfite; amplifying the nucleic acid in each reaction vessel;

cleaving the nucleic acids in each reaction vessel to produce fragments thereof;

obtaining a mass spectrum of the resulting fragments from one reaction vessel and another mass spectrum of the resulting fragements from another reaction vessel; whereby:

- 20 cytosine methylation is detected by identifying a difference in signals between the mass spectra.
 - 87. The method of claim 86, wherein: the step of amplifying is carried out in the presence of uracil; and the step of cleaving is effected by a uracil glycosylase.
- 25 88. A method for identifying a biological sample, comprising: generating a data set indicative of the composition of the biological sample;

denoising the data set to generate denoised data; deleting the baseline from the denoised data to generate an intermediate

30 data

15

set;

defining putative peaks for the biological sample;

using the putative peaks to generate a residual baseline;

removing the residual baseline from the intermediate data set to generate a corrected data set;

locating, responsive to removing the residual baseline, a probable peak in

5 the

corrected data set; and

identifying, using the located probable peak, the biological sample; wherein the generated biological sample data set comprises data from sense

10 strands and antisense strands of assay fragments.

89. The method according to claim 88, wherein identifying includes combining data from the sense strands and the antisense strands, and comparing the data against expected sense strand and antisense strand values, to identify the

15 biological

sample.

- 90. The method according to claim 88, wherein identifying includes deriving a peak probability for the probable peak, in accordance with whether the probable peak is from sense strand data or from antisense strand data.
- 20 91. The method according to claim 88, wherein identifying includes deriving a peak probability for the probable peak and applying an allelic penalty in response to a

ratio between a calculated area under the probable peak and a calculated expected average area under all peaks in the data set.

92. A method for identifying a biological sample, comprising: generating a data set indicative of the composition of the biological sample;

denoising the data set to generate denoised data;

deleting the baseline from the denoised data to generate an intermediate

30 data set:

25

20

30

defining putative peaks for the biological sample; using the putative peaks to generate a residual baseline;

removing the residual baseline from the intermediate data set to generate

5 corrected data set;

locating, responsive to removing the residual baseline, a probable peak in the corrected data set; and

identifying, using the located probable peak, the biological sample; wherein identifying includes deriving a peak probability for the probable peak and

applying an allelic penalty in response to a ratio between a calculated area under the

probable peak and a calculated expected average area under all peaks in the data set.

15 93. The method according to claim 92, wherein identifying includes comparing

data from probable peaks that did not receive an applied allelic penalty to determine their mass in accordance with oligonucleotide biological data.

- 94. The method according to claim 92, wherein the allelic penalty is not applied to probable peaks whose ratio of area under the peak to the expected area value is greater than 30%.
- 95. A method for detecting a polymorphism in a nucleic acid, comprising:

amplifying a region of the nucleic acid to produce an amplicon, wherein the resulting amplicon comprises one or more enzyme restriction sites;

contacting the amplicon with a restriction enzyme to produce fragments; obtaining a mass spectrum of the resulting fragments and analyzing signals in the mass spectrum by the method of claim 88; whereby:

the polymorphism is detected from the pattern of the signals.

96. A subcollection of samples from a target population, comprising: a plurality of samples, wherein the samples are selected from the group consisting of nucleic acids, fetal tissue, protein samples; and

20

a symbology on the containers containing the samples, wherein the symbology is representative of the source and/or history of each sample, wherein:

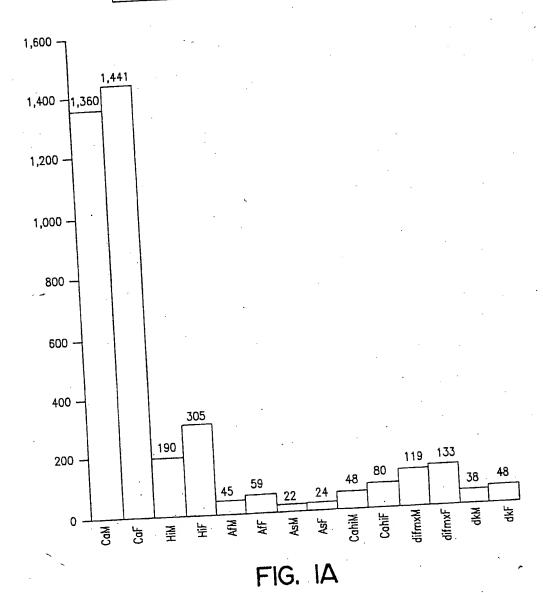
the target population is a healthy population that has not been selected for any disease state;

the collection comprises samples from the healthy population; and the subcollection is obtained by sorting the collection according to specified parameters.

- 97. The combination of claim 26, wherein the samples are selected selected from the group consisting of nucleic acids, fetal tissue, protein, tissue, body fluid, cell, seed, microbe, pathogen and reproductive tissue samples.
- 98. A combination, comprising the database of claim 8 and a mass spectrometer.
- 99. The combination of claim 98 that is an automated process line for analyzing biological samples.
 - 100. A system for high throughput processing of biological samples, comprising:
 - a database of claim 8, wherein the samples tested by the automated process line comprise samples from subjects in the database; and a mass spectrometry for analysis of biopolymers in the samples.

DNA Bank

		-	
			7010
	ء ـ	Samples	3912 l
Number	ΟI	2011bics	
Mullibo.			

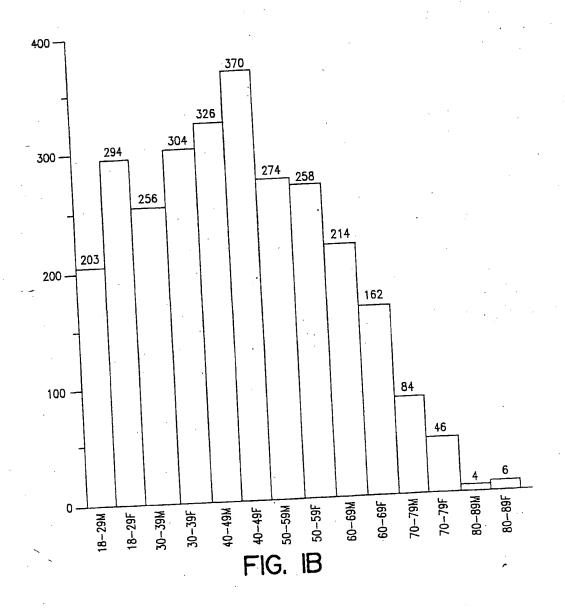


SUBSTITUTE SHEET (RULE 26)

2/51

Caucasians

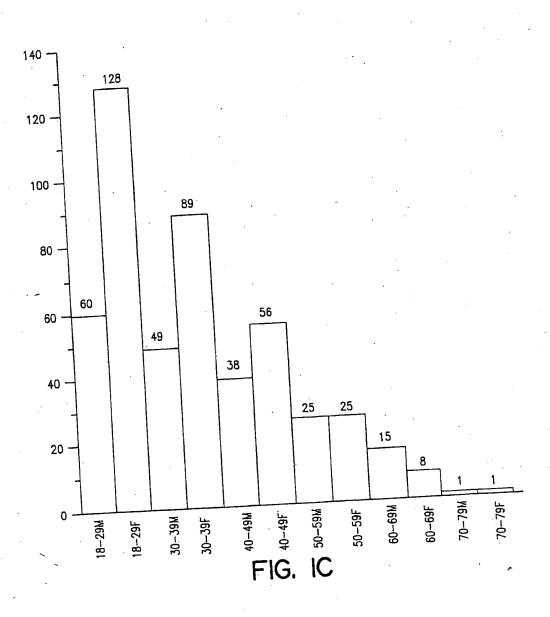
				2801	
ı	11	٦f	Samples	L 2001 1	í
١	Number	Οl	Samples		١
	110111				



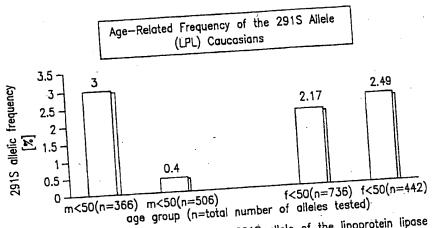
SUBSTITUTE SHEET (RULE 26)

3 / 51 Hispanics

			105
		Camples	495 I
Number	OΤ	Samples	
(NULLIDO)	•		

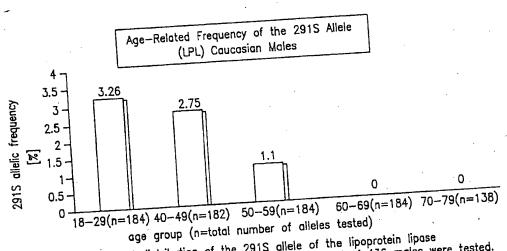


SUBSTITUTE SHEET (RULE 26)



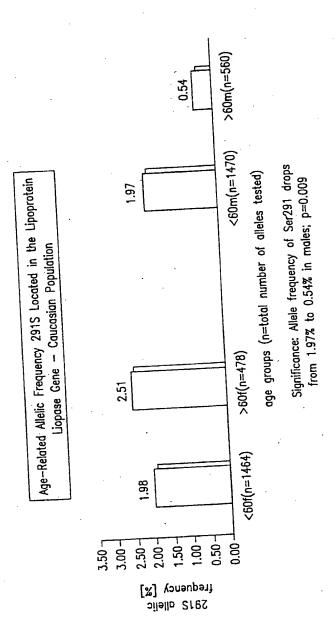
age— and sex—distribution of the 2915 allele of the lipoprotein lipase gene. A total of 436 males and 586 females were investigated.

FIG. 2A



Age— related distribution of the 291S allele of the lipoprotein lipase gene within the male Caucasian population. A total of 436 males were tested.

FIG. 2B



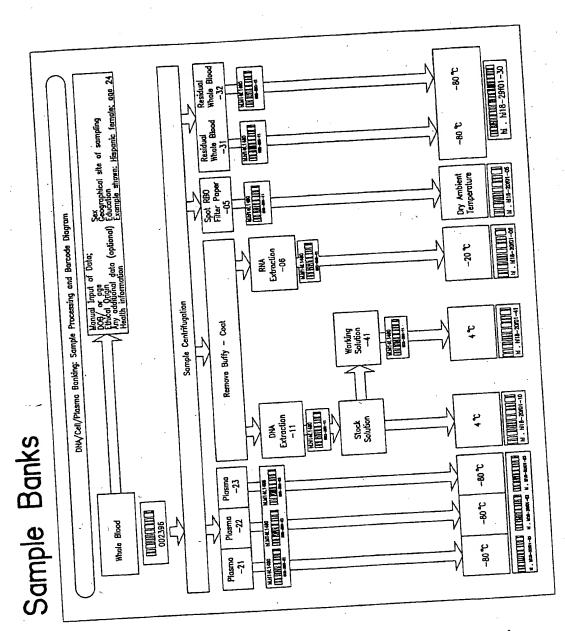
SUBSTITUTE SHEET (RULE 26)

6/51

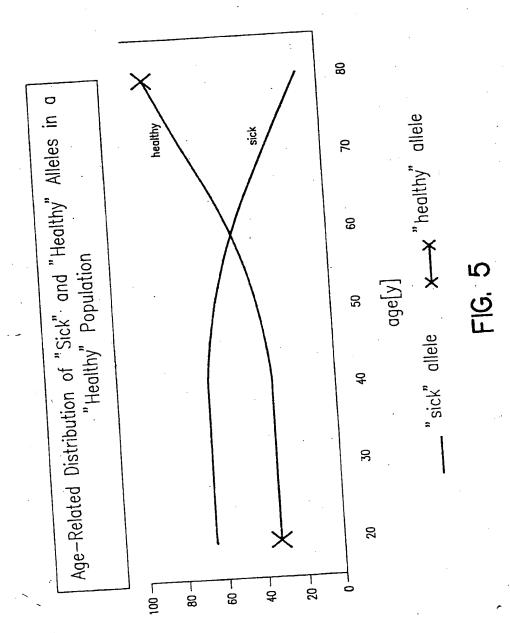
Questionnaire for Population—Based Sample Banking

Data Collection Form
Collection Information
Consent Form Signed Yes No Date of Collection (MM/DD/YY)/_/98 Time of Sample Collection(nearest hour in 24 hour clock format) Affix Barcode Here Initials of Data CollectorCollecting Agencyintactlostbroken (DO NOT COMPLETE: (For Date Entry Only)Sampleintactlostbroken
Donor information Date of Birth (MM/YY) / Years
Sex: Male Female How long have you lived there?
In which state do you live?
college graduate 4 J. 333.44
To the best of your knowledge what is the Ethnic Origin of your.
1
Courseign (please check specific geographic died ballow Nethodonde Norway Sweden, Switzerland, U.K.)
Northern Furnne (AUSTIQ, Delition, Internal Control
Southern Europe (Greece, Ruly, Spanny Pussin Yugoslavia)
Southern Europe (Greece, Italy, Spain) Czechostowakia, Hungary, Poland, Russia, Yugoslavia) Eastern Europe (Czechostowakia, Hungary, Poland, Russia, Yugoslavia) (Tzech Egypt Transland, Jordan, Syria, other Arab States)
☐ ☐ Eastern Europe (Czechoslovakia, Hungary, Pokina, Kassak, System States) ☐ ☐ Middle Eastern (Israel, Egypt, Iran, Iraq, Jordan, Syria, other Arab States)
and the state of t
Hieranic (glease check specific geographic died balou it thanks)
T a Verico
Central Amenca South Attended
- A.L. Duarto Rich filling Cullipodii
Cuba, Fuerto fileo, Ostanio Incorporation fileo, Ostanio
D D Japanese
🗆 🗈 Chinese 🛸
□ □ Korean
□ Vietnamese
other Asian
Other
□ □ Don't know Health information: Have you or has anyone in your immediate family(parents,brothers,sisters, or your children)
Health information: Have you or has anyone in your immediate family parameters.
had the following? Check all that apply
had the following? Check all that apply You Mother Father Sister Brother Child Disease:
Heart Disease Stroke or Arteriosclerosis
Concer (Specify type II knowi)
All times Disease of Dementio
to the market of Autoimmune Disease
Nervous System Disease like Mulaple Scial State
Other (please specify)
Additional health information details you would like to provide:
Additional nearly information details year

FIG. 3

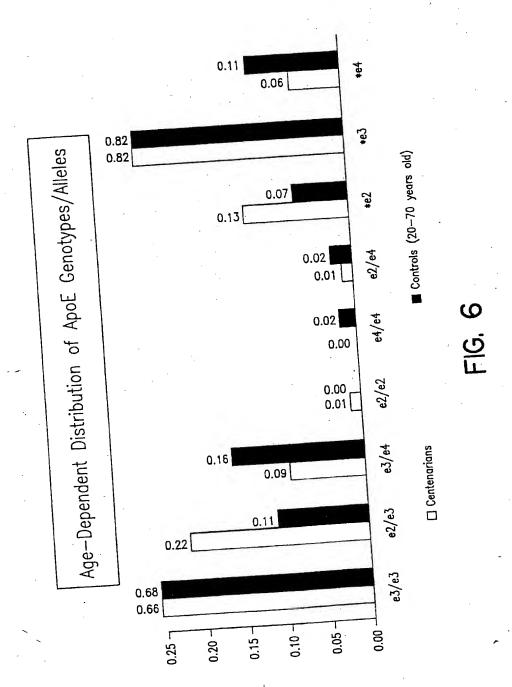


-1G, 4

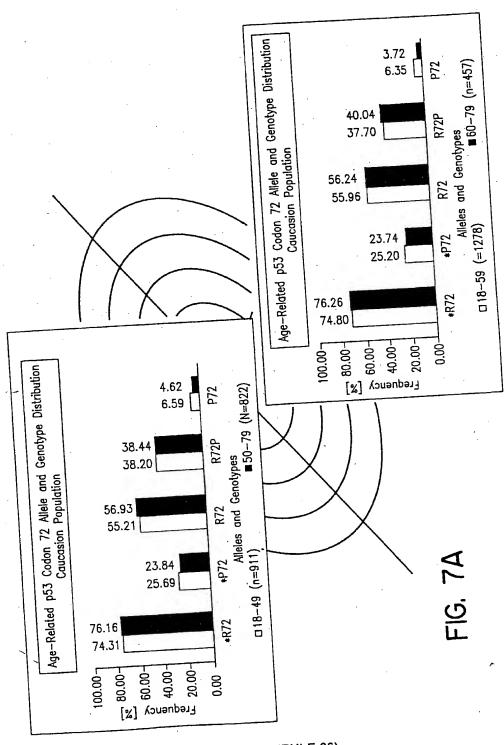


allelic frequency [%]

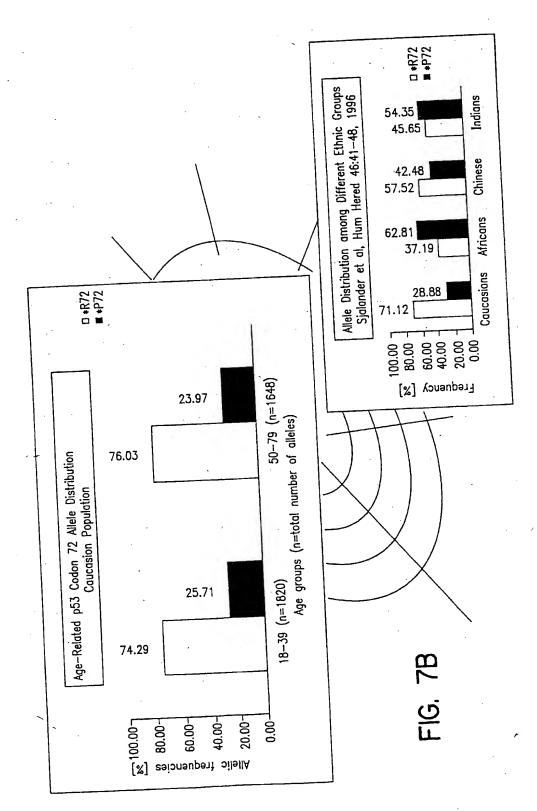
SUBSTITUTE SHEET (RULE 26)



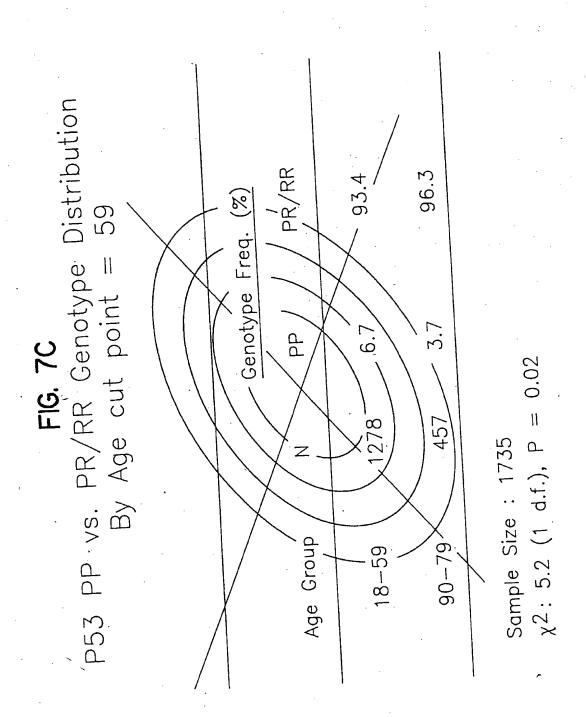
SUBSTITUTE SHEET (RULE 26)



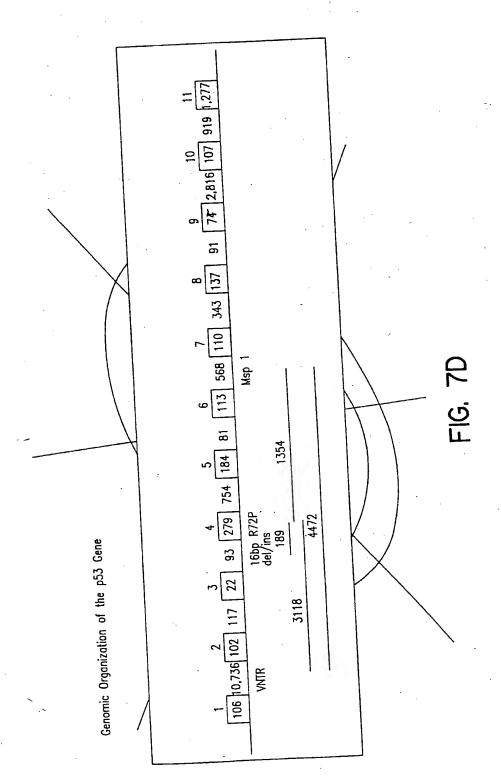
SUBSTITUTE SHEET (RULE 26)



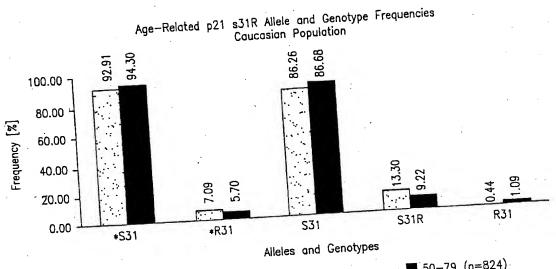
SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)

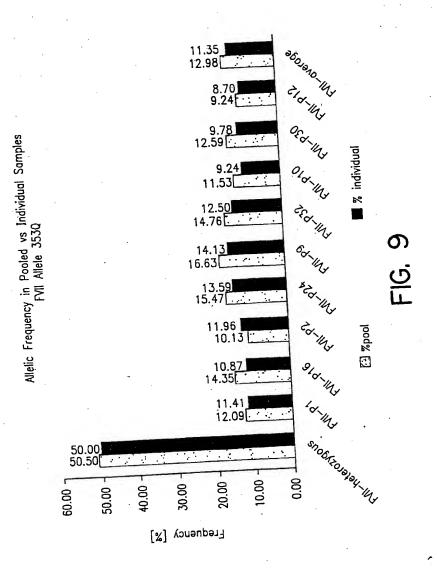


☐ 18-49 (n=910)

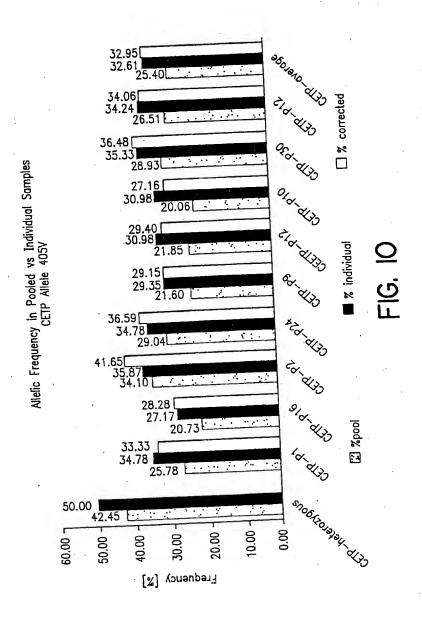
■ 50-79 (n=824)

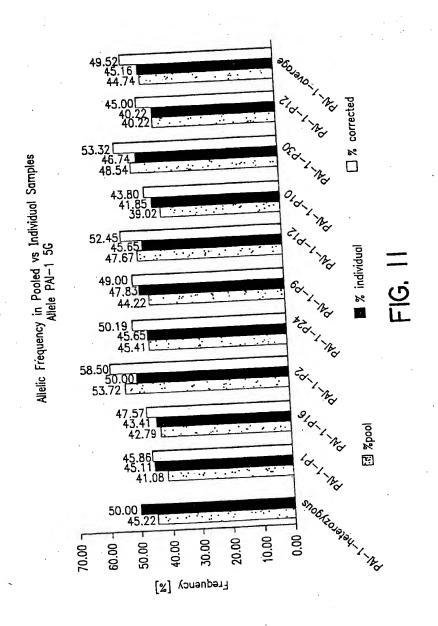
Significance: Genotype frequencyof SR heterozygous drops from 13.3% to 9.2%; p=0.009

FIG. 8



SUBSTITUTE SHEET (RULE 26)

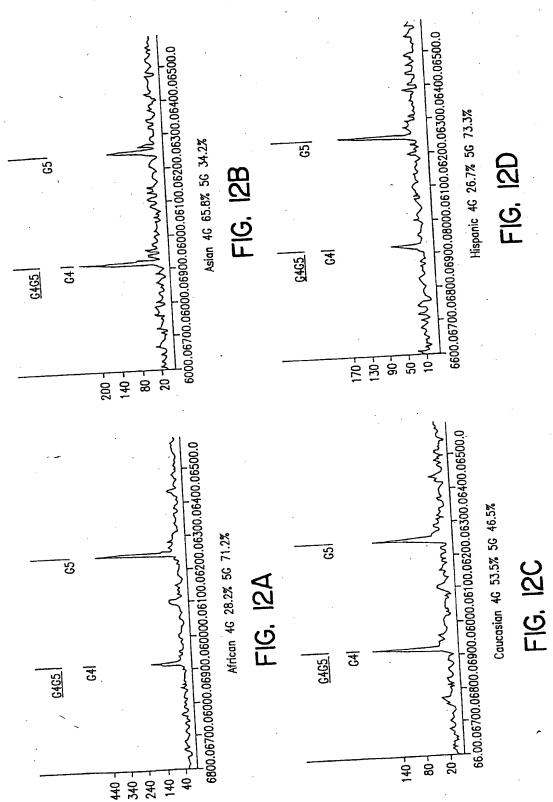




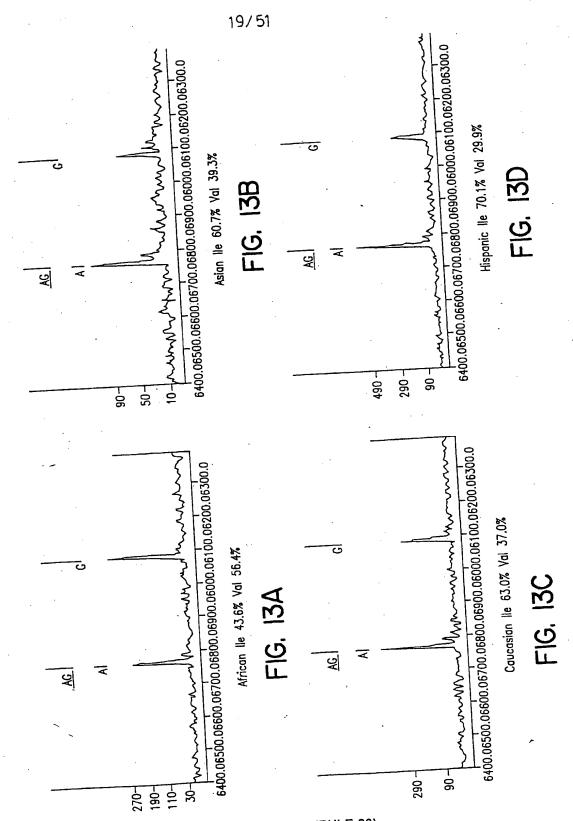
SUBSTITUTE SHEET (RULE 26)



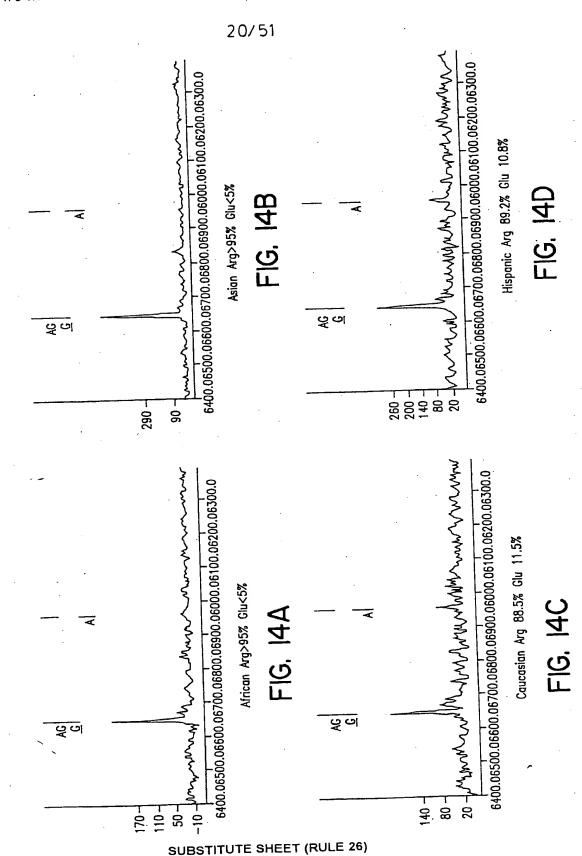
18/51

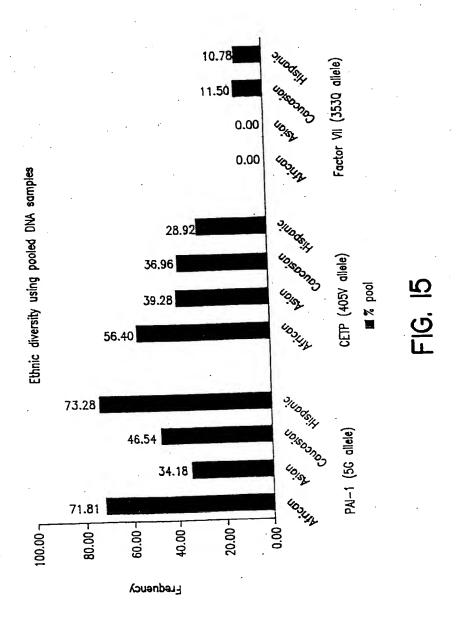


SUBSTITUTE SHEET (RULE 26)

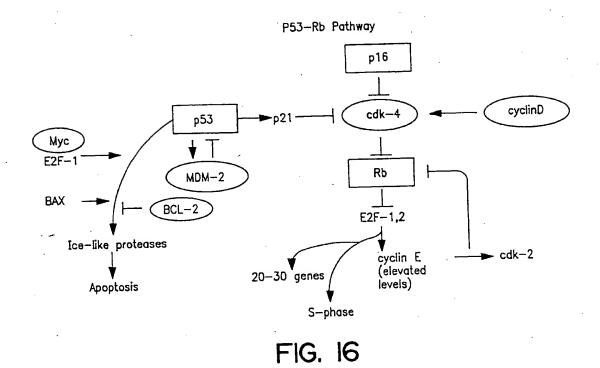


SUBSTITUTE SHEET (RULE 26)





SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)

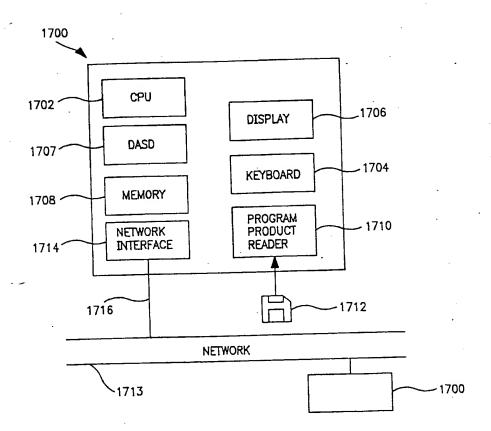


FIG. 17

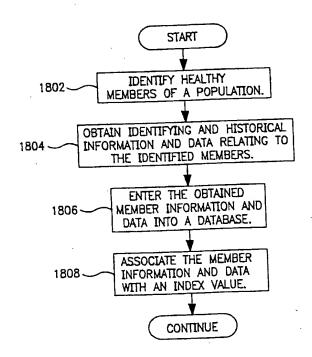
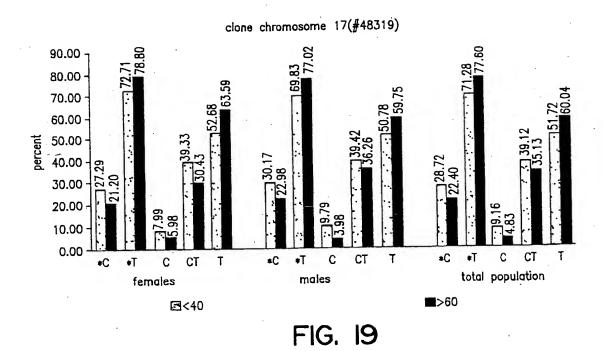


FIG. 18



SUBSTITUTE SHEET (RULE 26)

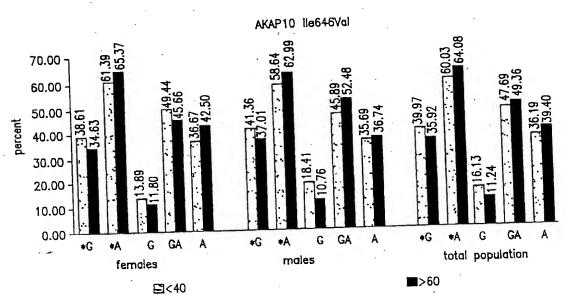


FIG. 20

methionine sulfoxide reductase A (#63306)

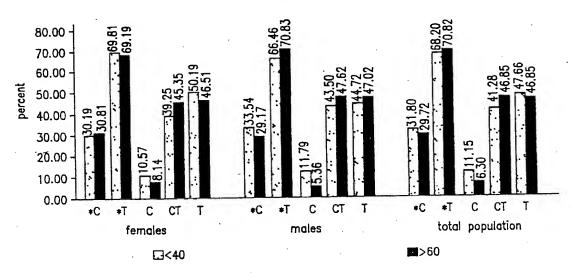


FIG. 21

Collec	tion Inform	nation								•	
	ent Form S 7 Yes	Signed No		Time of Sampl	• [Initials	Initials of Da	ta Collector			
JAN FEB MAR APR MAY JUN JUL AUG SEP OCT				(neorest hour, 24 hour cloc formet) 1888						BAR CODE	
DEC							 -				
	r Informat					····					
JAN FEB MAF APR MAY JUN JUL AUG SEF OCT			Sex: Male Female			Anat Physical activity do you do no regular base. Running Swimming Biking Gymnastics Other None		n? How many times have you been pregnant?	time	mony give BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB	
	To the	heat of	vous knowleds	je, what is the	Ethnic Origin	of your:					
	E 0 0 0 0 0 0		Coucasian () Northern Eur Southern Eur	alease mark spe rope (Austria, Di rope (Greece, It	ecific geograp enmark, Finla aly, Spain, Tu ekia, Hungary	nhic area below nd, France, Gen	many, Netheric L Yuqoslavia)	anda, Norway, Sw ates)	eden, Swi	itzerland, UK)	
			African-Ame	rican							
n which itate do ou live		0000	Mexico Central Ame	rico, South Ame o Rico, other Co	rica	ic area belaw if	known)				
	000000	000000	Asian (pleas Japanese Chinese Korean Vietnamese Filipino	e mark specific	geographic	area below if kr	nown)				
			Native Amer	ican							
988	<u></u> =	0.0	Other Don't know			·					}
	How long have you lived there?		⇒high schoo equlvalen some colle ⇒college gra	in school? 8th grade n,or11th grade I graduate or	☐ Yes ☐ No If Yes at what age	30-39 □ 40-49 □ 50-59 □ 60-69 □ 70-79 □ 80-89 ≥ 90	Heart Disease Cancer Stroke Accident Suicide	If Yes at what age?		of Death Falt Heart Dise Cancer Strake Accident Suicide	

29/51

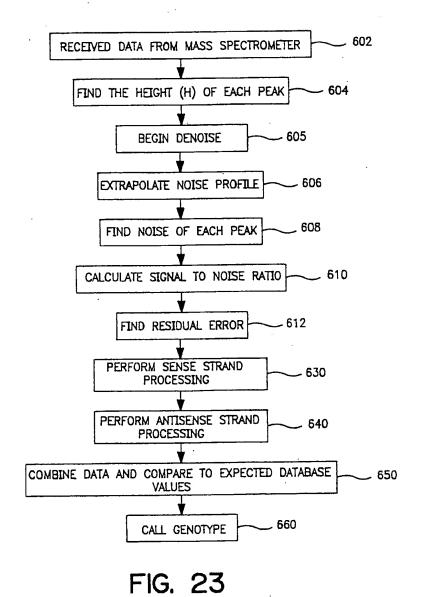
Have you ever smoked? Yes No Have you been hospitalized in the past 5 years for mo then 6 days at a time? Years Yes No Wanny times? Years Yes No wanny times? Yes No No Have you been hospitalized in the past 5 years for mo then 6 days at a time? Years Yes No No Have you been hospitalized in the past 5 years for mo then 6 days at a time? Years Yes No	Chronic Disorder Accident Chronic Disorder, including infection and thrombosis Chronic Disorder Accident Other: 3) Weeks: Acute disorder, including infection and thrombosis Chronic Disorder Accident Chronic Disorder Acute disorder, including infection and thrombosis Chronic Disorder Accident
Mark all that apply! Disease Heart Disease, including arteriosclerosis Stroke Hypertension Blood clats Diabetes, insulin dependent Diabetes, not insulin-dependent (diet controlled) Cancer: Lung&Bronchus Breasts Prostate Colon&Recturn Skin Lymphoma&Leukemia	You Mother Fother Sister Brother Child O O O O O O O O O O O O O O O O O O O
Other, please specify below: Alzheimer's Disease Epilepsy Schizophrenia	000
Bipolar disorder (manic depression) Major depression Chronic Inflammatory or Autoimmune Disease including Multiple Sclerosis and Rheumatoid Arthritis Emphysema Asthma Other, please specify below:	00 000 00 000 00 000 00 000
Do you take prescription drugs on a regular basis? If yes, please specify below: Have you ever donated blood before? Yes No Additional	☐ Yes ☐ No health information details you would like to provide:
If yes, how many times: Number of Times COCOCCO COCCOCCO COCCOCCO COCCOCCO COCCOC	FOR OFFICE USE ONLY
Do you drink any kind of alcohalic beverage? Never Less than 3 times per week Daily	යා යා යා යා යා යා

FIG. 22B

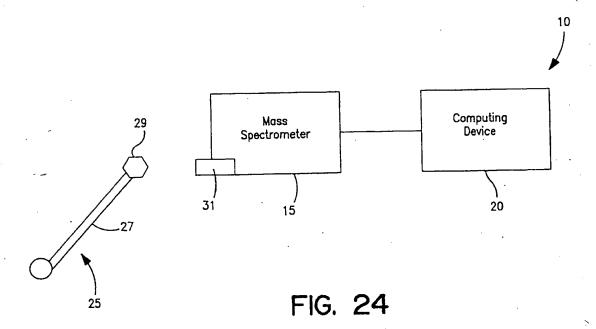
_								
Callection Info	rmation							
Consent Form	Signed No	Time of Sample Collection	I nitials	Initia	ts of Data Co	llector	•	
Month JAN FEB MAR JUN JUN JUN JUN JUN JUN JUN JU	Collection oy Year 2 0 0 0 100 0000000000000000000000000	(negrest hour in 24hour clock format) IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	######################################		(DO NOT COME for data entry ample: Intact Lost Broken	Volume (m)		•
DEC 🗀								
Donor Inform	nation						_	
JAN DO FEB MAR DO GAN D	Year	Ft. Inches Idle Image: Inches Ima	on a re	do you do vi egular basis? nning imming ding mnastics her	re you a egetarian? Yes No	If female: How many times have pregnant? It is	How many times did myou give thinth?	
In which sto	nte							
do you live?	To the heet of your	knowledge, what is the	Ethnic Origin of	your:				
	Egither Mother Co. See See See See See See See See See Se	ucasion (please mark sp rthern Europe (Austria, l uthern Europe (Greece, stern Europe(Czechoslow ddle Eastern (Israel, Egy her	pecific geographic Denmark, Finland Italy, Spain, Turk	c grea below i , France, Germ ey) pland Russia.	Yugoslavia)		en, Switzerland, Ul	‹›
		in't know						
How long	Years [have you smoking? Years Did smo	mar ago you quit sking?	Years do/	Years		f yes, for now long?	
have you there?	(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)		Yes No	884 884 886 888 888 888 888 888	100 100 100 100 100 100 100 100 100 100	No No	Contin Co	
			226					

What is your highest grade <u>Mother</u> Deceased? Cause of Death Mathe	r: <u>Father</u> Dec	eqsed?	Cause of	Death Fat	her:		
What is your highest grade you completed in school? less then 8th grade Yes No Bth,9th,10th,or11th grade No high school graduate or equivalency some callege, 2yr degree callege graduate,4yr degree post graduate education or degree 900 Mother Deceased? Cause of Death Mother Deceased. Cause	Yes No If Yes at what age?	≤ 25 30-3 40-4 50-5 60-6	Head Strong Access Suice 9 Coh	rt Disease cer oke ident cide			
Health Information		hildron)	had the	following?	*		
Have you or has anyone in your immediate family (parents,brothers,sist Mark all that apply! Disease Heart Disease Stroke Hypertension Blood clots Diabetes, insulin dependent Diabetes, not insulin-dependent Cancer: Lung&Bronchus Breasts Prostate Calon&Rectum Skin Lymphama&Leukemia Other, please specify below:		**************************************	Father 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0			30000000000000000	,
Alzheimer's Disease Epilepsy Schizophrenia	000	000	000	000	000	000	
Bipolar disorder (manic depression) Major depression Chronic Inflammatory or Autoimmune Disease including Multiple Sciences and Rheumataid Arthritis Emphysema Asthma Other, please specify below:	00.000	00 000	00 000	00 000	000 00	000 00	
Do you take prescription drugs on a regular basis? If yes, please specify below:		Yes ⊏	II P	ove you o lood before yes, how nany times	s? Nur	mber of Ti	imes
Have you been hospitalized in the past 5 years for more then 6 days at a time? Yes No Chronic Disorder For each hospitalization (if not the same) how long did you stay and for what reason? 1) Weeks: Chronic Disorder Acute disorder, includic Chronic Disorder	infection of the control of the cont	and thro	mbosis		-		86896888
Do you drink any kind of alcoholic beverage? Never					FOR CUSE	OFFICE ONLY	
					E	භ	

FIG. 22D



SUBSTITUTE SHEET (RULE 26)



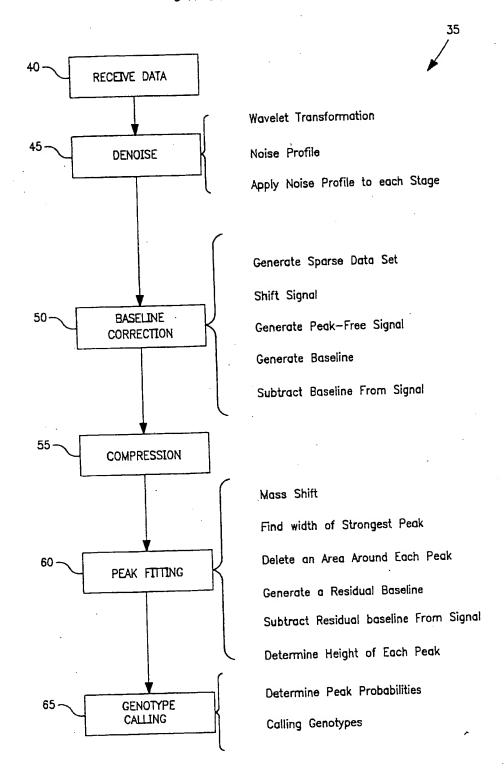


FIG. 25
SUBSTITUTE SHEET (RULE 26)

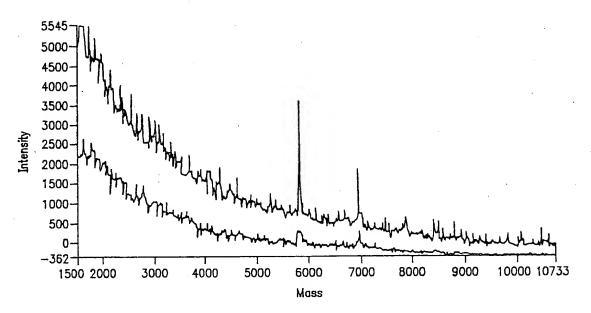
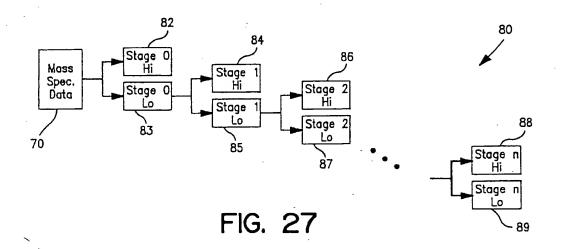
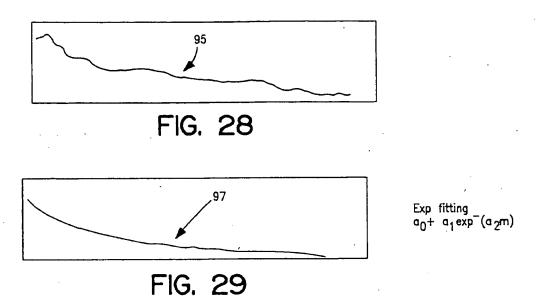


FIG. 26





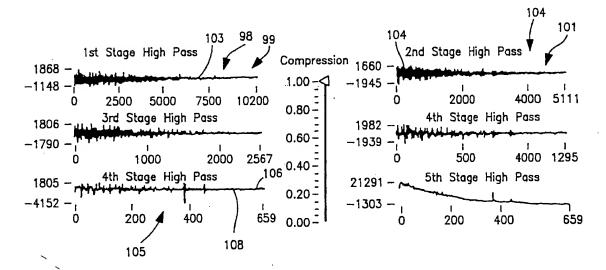
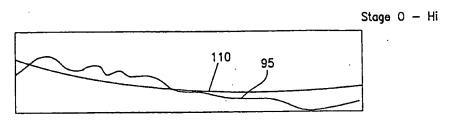
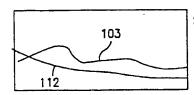


FIG. 30

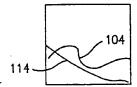


Threshold 0=4XNoiseProfile



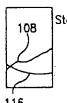
Stage 1 - Hi

Threshold 1=2XNoiseProfile



Stage 2 - Hi

Threshold 2=1XNoiseProfile



Stage n — Hi

Threshold n=(1/2 n-2)XNoiseProfile

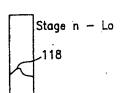
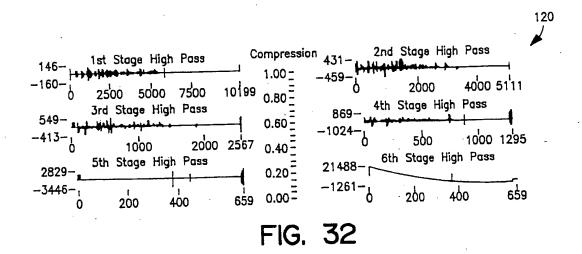


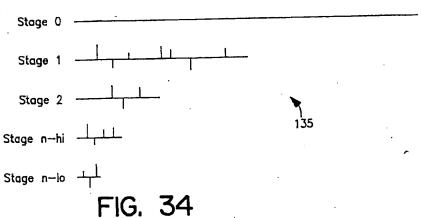
FIG. 31



Signal (t)=
$$\frac{\text{(Start 0(t) + Start 1(t) + Start 2(t)... + Start 23 (t))}}{24}$$

SHIFT SIGNAL TO ACCOUNT FOR VARIATIONS DUE TO STARTING POINT

FIG. 33



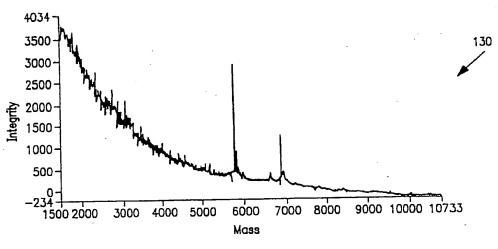


FIG. 35

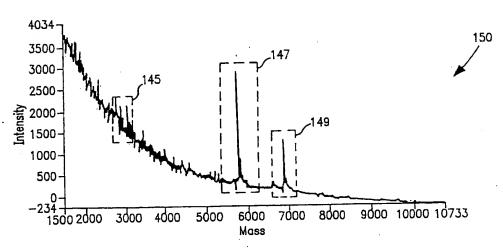


FIG. 13-TAKE A MOVING AVERAGE, REMOVE SECTIONS EXCEEDING A THRESHOLD

FIG. 36

40/51

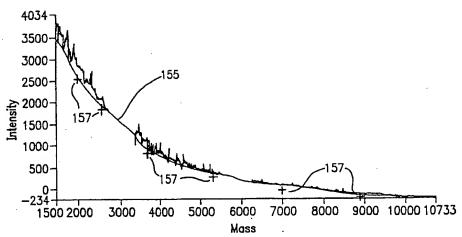


FIG. 37

FIND MINIMA IN REMAINING SIGNALS AND CONNECT TO FORM A PEAK FREE SIGNAL

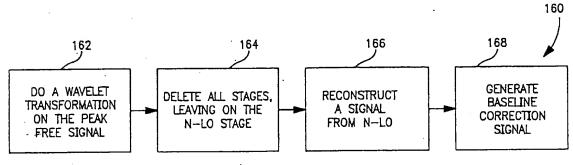


FIG. 38

GENRATE BASLELINE CORRECTION

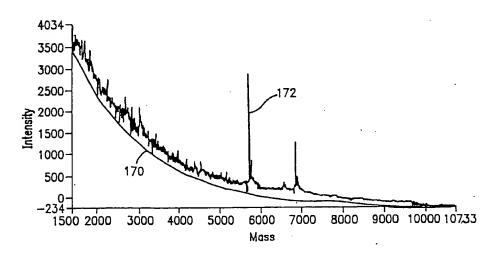
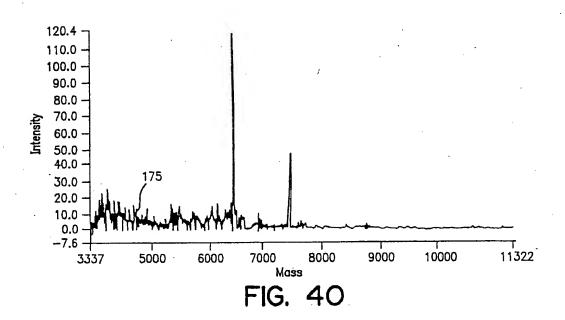


FIG. 39



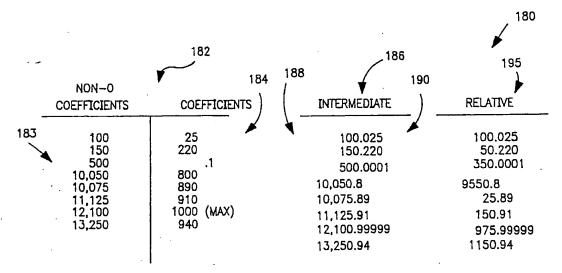


FIG. 41

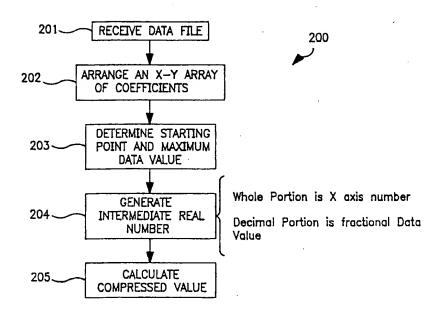


FIG. 42

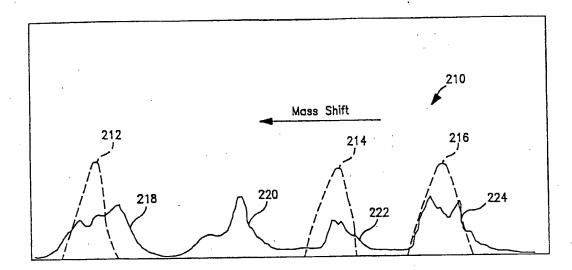
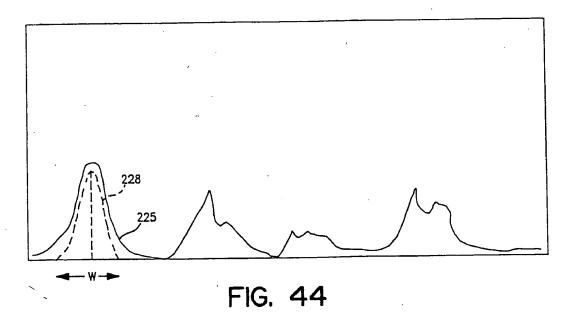
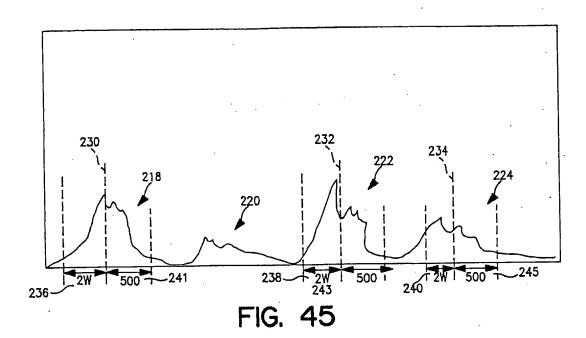


FIG. 43



SUBSTITUTE SHEET (RULE 26)



245

FIG. 46

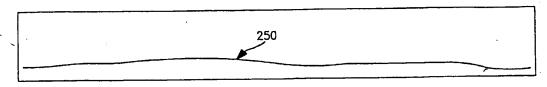


FIG. 47

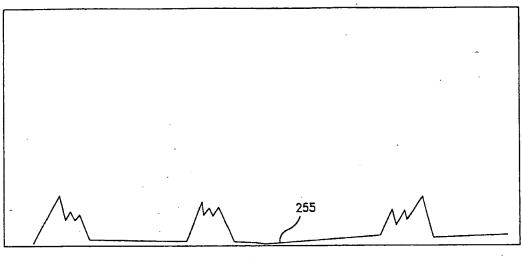


FIG. 48

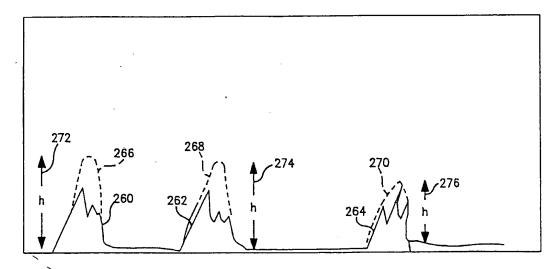
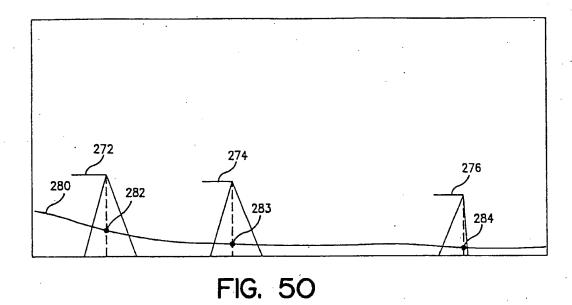
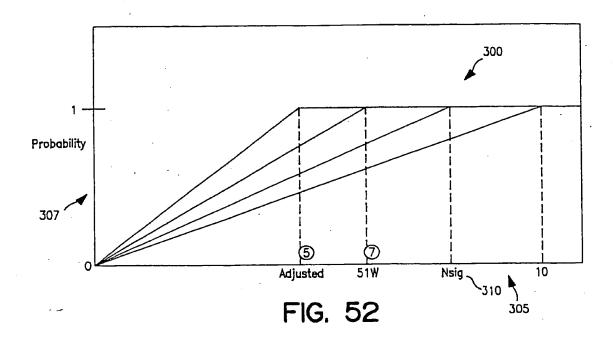


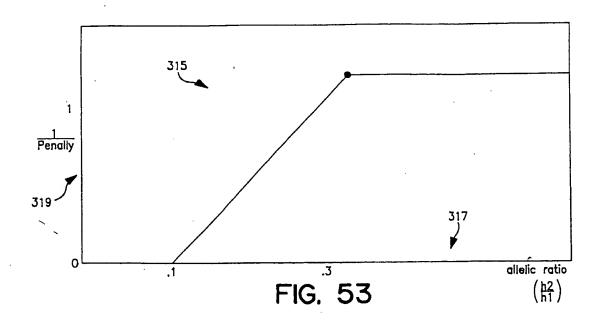
FIG. 49



293(G) 290(n) FIG. 5I

SUBSTITUTE SHEET (RULE 26)





SUBSTITUTE SHEET (RULE 26)

48/51.

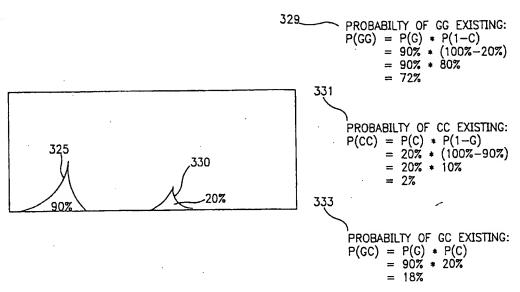
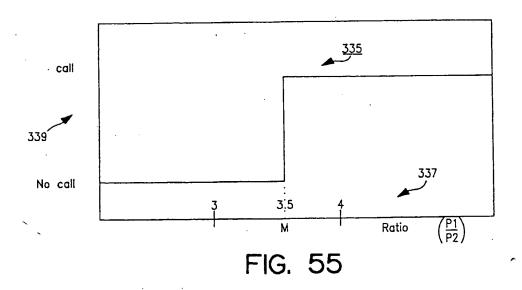


FIG. 54



SUBSTITUTE SHEET (RULE 26)

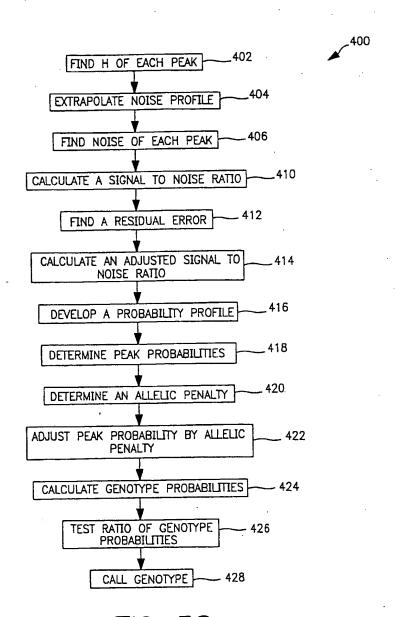


FIG. 56

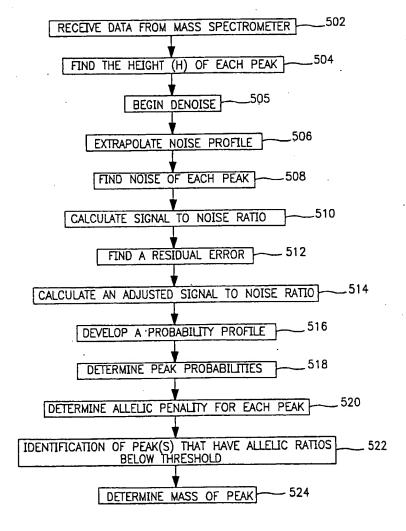
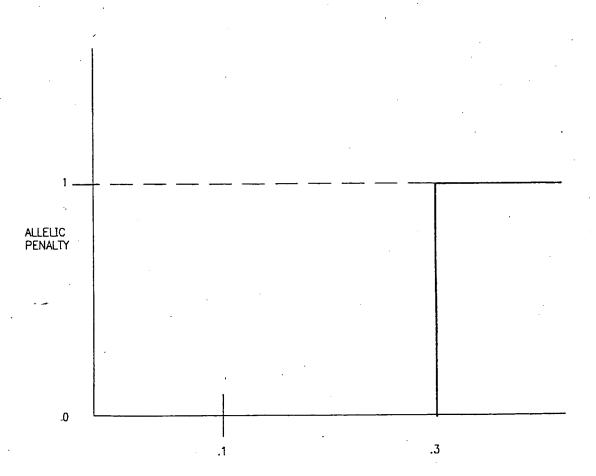


FIG. 57

51/51



RATIO OF AREA UNDER PEAK

FIG. 58

SEQUENCE LISTING

```
<110> SEQUENOM
        Braun et al.
<120> METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING
POLYMORPHIC GENETIC MARKERS
<130> 24736-2033PC
<140> Not Yet Assigned
<141> 2000-10-13
<150> 60/217,658
<151> 2000-07-10
<150> 60/159,176
<151> 1999-10-13
<150> 60/217,251
<151> 2000-07-10
<150> 09/663,968
<151> 2000-09-19
<160> 118
<170> FastSEQ for Windows Version 4.0
<210> 1
<211> 361
<212> DNA
<213> Homo Sapien
ctgaggacct ggtcctctga ctgctctttt cacccatcta cagtccccct tgccgtccca agcaatggat gatttgatgc tgtccccgga cgatattgaa caatggttca ctgaagaccc
                                                                                         120
aggtecagat gaageteeca gaatgecaga ggetgeteec egegtggeec etgeaceage ageteetaca eeggeggeec etgeaceage ecceteetgg eccetgteat ettetgteec
                                                                                         180
                                                                                         240
ttcccagaaa acctaccagg gcagctacgg tttccgtctg ggcttcttgc attctgggac agccaagtct gtgacttgca cggtcagttg ccctgagggg ctggcttcca tgagacttca
                                                                                         300
                                                                                         360
                                                                                         361
<210> 2
<211> 44
<212> DNA
<213> Artificial Sequence
<223> Oligonucleotide Primer
                                                                                          44
cccagtcacg acgttgtaaa acgctgagga cctggtcctc tgac
<210> 3
<211> 42
<212> DNA
<213> Artificial Sequence
<220>
<223> Oligonucleotide Primer
                                                                                          42
ageggatàac aattteacae aggttgaagt eteatggaag ee
<210> 4
<211> 17
<212> DNA
```

<213>	Artificial Sequence	
<220> <223>	Probe	
<400>	4	
gccaga	agget getecee	17
<210>	5	
<211>	17	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Probe	
<400>		
		17
gccaga	agget geteece	1,
<210>		
<211>		
<212>	DNA	
<213>	Artificial Sequence	
<220>	·	
<223>	Probe	
<400>		1.0
gccaga	agget geteceege	19
<210>	7	
<211>	18	
<212>		
<213>	Artificial Sequence	
<220>		
<223>	Probe	
	_	
<400>		18
gccaga	agget geteecee	10
<210>	8	
<211>	161	
<212>		
<213>	Homo Sapien	
<400>	8	
	caga acccatgogg cagcaagged tgddgccgdd tettoggddd agtggadagd	60
	octga geogegaetg tgatgegeta atggeggget geatecagga ggeoegtgag	120
	gaact togactttgt cacogagaca coactggagg g	161
<210>	9	
<211>		
<212>		
	Artificial Sequence	
<220>	Oligonyal optido Drimor	
<223>	Oligonucleotide Primer	
<400>	9	

cccagtcacg acgttgtaaa acggtccgtc agaacccatg cgg	43
<210> 10 <211> 44 <212> DNA <213> Artificial Sequence	
<220> <223> Oligonucleotide Primer	
<400> 10 agcggataac aatttcacac aggctccagt ggtgtctcgg tgac	44
<210> 11 <211> 15 <212> DNA <213> Artificial Sequence	
<220> <223> Oligonucleotide Primer	
<400> 11 cagcgagcag ctgag	15
<210> 12 <211> 15 <212> DNA <213> Artificial Sequence	-
<220> <223> Probe	
<400> 12 cagcgagcag ctgag	15
<210> 13 <211> 16 <212> DNA <213> Artificial Sequence	
<220> <223> Probe	
<400> 13 cagcgagcag ctgagc	16
<210> 14 <211> 17 <212> DNA <213> Artificial Sequence	
<220> <223> Probe	
<400> 14 cagcgagcag ctgagac	17
<210> \15	

```
<213> Homo Sapien
<400> 15
gegetecatt catetettea tegactetet gttgaatgaa gaaaateeaa gtaaggeeta caggtgeagt tecaaggaag cetttgagaa agggetetge ttgagttgta gaaagaaceg etgeaacaat etgggetatg agateaataa agteagagee aaaagaagea geaaaatgta cetgaagaet egtteteaga tgeee
                                                                                             60
                                                                                            120
                                                                                            180
                                                                                            205
<210> 16
<211> 42
<212> DNA
<213> Artificial Sequence
<223> Oligonucleotide Primers
<400> 16
                                                                                             42
cccagtcacg acgttgtaaa acggcgctcc attcatctct tc
<210> 17 <211> 42
<212> DNA
<213> Artificial Sequence
<220>
<223> Oligonucleotide Primer
                                                                                             42
agcggataac aatttcacac agggggcatc tgagaacgag tc
<210> 18
<211> 20
<212> DNA
<213> Artificial Sequence
<223> Oligonucleotide Primer
<400> 18
caatctgggc tatgagatca
                                                                                             20
<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Probe
<400> 19
                                                                                             20
caatctgggc tatgagatca
<210> 20
<211> 21
<212> DNA
<213> Artificial Sequence
<220> \
<223> Probe
```

```
<400> 20
caatctgggc tatgagatca a
                                                                             21
<210> 21
<211> 22
 <212> DNA
 <213> Artificial Sequence
<220>
<223> Probe
<400> 21
caatctgggc tatgagatca gt
                                                                                20
<210> 22
<211> 60
<212> DNA
<213> Homo Sapien
<220>
<223> Probe
<400> 22
gtgccggcta ctcggatggc agcaaggact cctgcaaggg ggacagtgga ggcccacatg
                                                                             60
<210> 23
<211> 60
<212> DNA
<213> Homo sapien
<400> 23
ccacccacta ccggggcacg tggtacctga cgggcatcgt cagctggggc cagggctgcg
                                                                             60
<210> 24
<211> 42
<212> DNA
<213> Artificial Sequence
<220>
<223> Oligonucleotide primer
<400> 24
cccagtcacg acgttgtaaa acgatggcag caaggactcc tg
                                                                             42
<210> 25
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Oligonucleotide primer
<400> 25
cacatgccac ccactacc
                                                                             18
<210> 26
<211> 43
<212> DNA
<213> Artificial Sequence
```

<220> <223> Oligonucleotide primer	
<400> 26 agggataac aatttcacac aggtgacgat gcccgtcagg tac	43
<210> 27 <211> 15 <212> DNA <213> Artificial Sequence	
<220> <223> Probe	
<400> 27 atgccaccca ctacc	15
<210> 28 <211> 19 <212> DNA <213> Artificial Sequence	
<220> <223> Probe	
<400> 28 cacatgocac coactacog	19
<210> 29 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> Probe	
<400> 29 cacatgccac ccactaccag	20
<210> 30 <211> 23 <212> DNA <213> Artificial Sequence	
<220> <223> Probe	
<400> 30 agcggataac aatttcacac agg	23
<210> 31 <211> 2363 <212> DNA <213> Homo Sapien	
<220> <221> CDS <222> (138)(2126) <223> AKAP-10	

7/122

<300> <308> GenBank AF037439 <309> 1997-12-21 <400> 31 60 geggettgtt gataatatgg eggetggage tgeetgggea teeegaggag geggtgggge ccactccogg aagaagggtc cettttcgcg ctagtgcagc ggcccctctg gacccggaag 120 teegggeegg ttgetga atg agg gga gee ggg eee tee eeg ege eag tee Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser 170 ccc cgc acc ctc cgt ccc gac ccg ggc ccc gcc atg tcc ttc ttc cgg Pro Arg Thr Leu Arg Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg 218 cgg aaa gtg aaa ggc aaa gaa caa gag aag acc tca gat gtg aag tcc 266 Arg Lys Val Lys Gly Lys Glu Glu Lys Thr Ser Asp Val Lys Ser att aaa gct tca ata tcc gta cat tcc cca caa aaa agc act aaa aat Ile Lys Ala Ser Ile Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn 314 cat gcc ttg ctg gag gct gca gga cca agt cat gtt gca atc aat gcc His Ala Leu Leu Glu Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala 362 att tot god aac atg gad tod ttt toa agt agd agg aca god aca ott Ile Ser Ala Asn Met Asp Ser Phe Ser Ser Ser Arg Thr Ala Thr Leu 410 85 aag aag cag cca agc cac atg gag gct gct cat ttt ggt gac ctg ggc Lys Lys Gln Pro Ser His Met Glu Ala Ala His Phe Gly Asp Leu Gly 458 aga tet tgt etg gae tae eag aet eaa gag aee aaa tea age ett tet 506 Arg Ser Cys Leu Asp Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser aag acc ctt gaa caa gtc ttg cac gac act att gtc ctc cct tac ttc Lys Thr Leu Glu Gln Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe 554 att caa ttc atg gaa ctt cgg cga atg gag cat ttg gtg aaa ttt tgg Ile Gln Phe Met Glu Leu Arg Arg Met Glu His Leu Val Lys Phe Trp 602 tta gag gct gaa agt ttt cat tca aca act tgg tcg cga ata aga gca Leu Glu Ala Glu Ser Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala 650 160 165 cac agt cta aac aca atg aag cag agc tca ctg gct gag cct gtc tct ${\tt His}$ Ser Leu ${\tt Asn}$ Thr Met Lys ${\tt Gln}$ Ser Ser Leu ${\tt Ala}$ Glu ${\tt Pro}$ Val ${\tt Ser}$ 698 180 cca tct aaa aag cat gaa act aca gcg tct ttt tta act gat tct ctt Pro Ser Lys Lys His Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu 746 195 gat aag aga ttg gag gat tot ggc toa gca cag ttg ttt atg act cat 794

	Asp	Lys 205	Arg	Leu	Glu	Asp	Ser 210	Gly	Ser	Ala	Gln	Leu 215	Phe	Met	Thr	His	
							aat Asn										842
							tgt Cys										890
							cac His										938
							gcc Ala										986
							cta Leu 290										1034
	aat Asn 300	Thr	ttt Phe	acc Thr	aaa Lys	tat Tyr 305	ata Ile	tct Ser	cca Pro	gat Asp	gct Ala 310	gct Ala	aaa Lys	cca Pro	ata Ile	cca Pro 315	1082
							aat Asn										1130
							aac Asn										1178
							gag Glu										1226
	cat His	ttc Phe 365	tgt Cys	aaa Lys	tac Tyr	cag Gln	att Ile 370	gaa Glu	gtg Val	ctg Leu	acc Thr	agt Ser 375	gga Gly	act Thr	gtt Val	tac Tyr	1274
	ctg Leu 380	gct Ala	gac Asp	att Ile	ctc Leu	ttc Phe 385	tgt Cys	gag Glu	tca Ser	gcc Ala	ctc Leu 390	ttt Phe	tat Tyr	ttc Phe	tct Ser	gag Glu 395	1322
	tac Tyr	atg Met	gaa Glu	aaa Lys	gag Glu 400	gat Asp	gca Ala	gtg Val	aat Asn	atc Ile 405	tta Leu	caa Gln	ttc Phe	tgg Trp	ttg Leu 410	gca Ala	1370
	gca Ala	gat Asp	aac Asn	ttc Phe 415	cag Gln	tct Ser	cag Gln	ctt Leu	gct Ala 420	gcc Ala	aaa Lys	aag Lys	Gly 999	caa Gln 425	tat Tyr	gat Asp	1418
		Gln					gat Asp										1466
•	tcc	ctc	caa	gcc	aca	cat	cct	ctt	gga	ttt	gat	gat	gtt	gta	cga	tta	1514

Ser	Leu 445		Ala	Thr	His	Pro 450	Leu	Gly	Phe	Asp	Asp 455	Val	Val	Arg	Leu	
					atc Ile 465											1562
					cgt Arg											1610
					tcc Ser											1658
					cga Arg											1706
ccg Pro	act Thr 525	gct Ala	cct Pro	ggc Gly	tct Ser	gtt Val 530	ggc	cct Pro	cct Pro	gat Asp	gag Glu 535	tct Ser	cac His	cca Pro	gl ^à aaa	1754
					gcg Ala 545											1802
aaa Lys	ata Ile	ctg Leu	aaa Lys	aat Asn 560	ttt Phe	gat Asp	gaa Glu	gcg Ala	ata Ile 565	att Ile	gtg Val	gat Asp	gcg Ala	gca Ala 570	agt Ser	1850
ctg Leu	gat Asp	cca Pro	gaa Glu 575	tct Ser	tta Leu	tat Tyr	caa Gln	cgg Arg 580	aca Thr	tat Tyr	gcc Ala	gjå aaa	aag Lys 585	atg Met	aca Thr	1898
					gac Asp											1946
					aaa Lys											1994
aag Lys 620	aaa Lys	tgg Trp	gtg Val	caa Gln	gga Gly 625	aat Asn	act Thr	gat Asp	gag Glu	gcc Ala 630	cag Gln	gaa Glu	gag Glu	cta Leu	gct Ala 635	2042
tgg Trp	aag Lys	att Ile	gct Ala	aaa Lys 640	atg Met	ata Ile	gtc Val	agt Ser	gac Asp 645	att Ile	atg Met	cag Gln	cag Gln	gct Ala 650	cag Gln	2090
					gag Glu						tga *	ctca	aaac	tt:		2136
caac gaat	acag 9999	icc a lag a	atga .caaț	aaac ccta	a go	acta cttc	tatt cacc	tct cta	gato	tgt agt	cact	gttg tgta	itt t	ccag	ittett gggaga iattgg	2196 2256 2316 2363

10/122

<210> 32

<211> 662 <212> PRT <213> Homo Sapien <400> 32 Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser Pro Arg Thr Leu Arg Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg Arg Lys Val Lys Gly Lys Glu Gln Glu Lys Thr Ser Asp Val Lys Ser Ile Lys Ala Ser Ile Ser Val His Ser Pro Gln Lys Ser Thr Lys Asn His Ala Leu Leu Glu Ala Ala Gly Pro Ser His Val Ala Ile Asn Ala Ile Ser Ala Asn Met Asp Ser Phe Ser Ser Ser Arg Thr Ala Thr Leu Lys Lys Gln Pro Ser His Met Glu Ala Ala His Phe Gly Asp Leu Gly Arg Ser Cys Leu Asp Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser Lys Thr Leu Glu Gln Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe Ile Gln Phe Met Glu Leu Arg Arg Met Glu His Leu Val Lys Phe Trp Leu Glu Ala Glu Ser Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala His Ser Leu Asn Thr Met Lys Gln Ser Ser Leu Ala Glu Pro Val Ser Pro Ser Lys Lys His Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu Asp Lys Arg Leu Glu Asp Ser Gly Ser Ala Gln Leu Phe Met Thr His Ser Glu Gly Ile Asp Leu Asn Asn Arg Thr Asn Ser Thr Gln Asn His Leu Leu Leu Ser Gln Glu Cys Asp Ser Ala His Ser Leu Arg Leu Glu Met Ala Arg Ala Gly. Thr His Gln Val Ser Met Glu Thr Gln Glu Ser Ser Thr Leu Thr Val Ala Ser Arg Asn Ser Pro Ala Ser Pro Leu Lys Glu Leu Ser Gly Lys Leu Met Lys Ser Ile Glu Gln Asp Ala Val Asn Thr Phe Thr Lys Tyr Ile Ser Pro Asp Ala Ala Lys Pro Ile Pro Ile Thr Glu Ala Met Arg Asn Asp Ile Ile Ala Arg Ile Cys Gly Glu Asp Gly Gln Val Asp Pro Asn Cys Phe Val Leu Ala Gln Ser Ile Val Phe Ser Ala Met Glu Gln Glu His Phe Ser Glu Phe Leu Arg Ser His His Phe Cys Lys Tyr Gln Ile Glu Val Leu Thr Ser Gly Thr Val Tyr Leu Ala Asp Ile Leu Phe Cys Glu Ser Ala Leu Phe Tyr Phe Ser Glu Tyr Met Glu Lys Glu Asp Ala Val Asn Ile Leu Gln Phe Trp Leu Ala Ala Asp Asn Phe Gln Ser Gln Leu Ala Ala Lys Lys Gly Gln Tyr Asp Gly Gln Glu Ala Gln

```
Asn Asp Ala Met Ile Leu Tyr Asp Lys Tyr Phe Ser Leu Gln Ala Thr
        435
                              440
His Pro Leu Gly Phe Asp Asp Val Val Arg Leu Glu Ile Glu Ser Asn
                          455
                                               460
Ile Cys Arg Glu Gly Gly Pro Leu Pro Asn Cys Phe Thr Thr Pro Leu
                     470
                                          475
Arg Gln Ala Trp Thr Thr Met Glu Lys Val Phe Leu Pro Gly Phe Leu
                                       490
                 485
Ser Ser Asn Leu Tyr Tyr Lys Tyr Leu Asn Asp Leu Ile His Ser Val
             500
                                  505
Arg Gly Asp Glu Phe Leu Gly Gly Asn Val Ser Pro Thr Ala Pro Gly
        515
                              520
                                                    525
Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly Ser Ser Asp Ser Ser
                          535
                                               540
Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile Lys Ile Leu Lys Asn
                     550
                                          555
Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser Leu Asp Pro Glu Ser
                 565
                                       570
                                                             575
Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr Phe Gly Arg Val Ser
                                  585
                                                        590
             580
Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro Glu Pro Asp Val Arg
                              600
                                                    605
        595
Lys Ser Lys Gly Ser Met Phe Ser Gln Ala Met Lys Lys Trp Val Gln
                          615
                                               620
    610
Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala Trp Lys Ile Ala Lys
                    630
                                          635
Met Ile Val Ser Asp Ile Met Gln Gln Ala Gln Tyr Asp Gln Pro Leu
                 645
                                    650
Glu Lys Ser Thr Lys Leu
<210> 33
<211> 2363
<212> DNA
<213> Homo Sapien
<220>
<221> CDS
<222> (138) ... (2126)
<223> AKAP-10-5
<221> allele
<222> 2073
<223> Single Nucleotide Polymorphism: A to G
<400> 33
                                                                            60
geggettgtt gataatatgg eggetggage tgeetgggea teeegaggag geggtgggge
ccactcccgg aagaagggtc ccttttcgcg ctagtgcagc ggcccctctg gacccggaag
tccgggccgg ttgctga atg agg gga gcc ggg ccc tcc ccg cgc cag tcc
Met Arg Gly Ala Gly Pro Ser Pro Arg Gln Ser
                                                                           120
                                                                           170
ecc ege acc etc egt ecc gae eeg gge ecc gee atg tec tte tte egg
                                                                           218
Pro Arg Thr Leu Arg Pro Asp Pro Gly Pro Ala Met Ser Phe Phe Arg
                                   20
                                                                           266
cgg aaa gtg aaa ggc aaa gaa caa gag aag acc tca gat gtg aag tcc
Arg Lys Val Lys Gly Lys Glu Glu Glu Lys Thr Ser Asp Val Lys Ser
```

					tcc Ser											314
					gct Ala 65											362
					gac Asp											410
					cac His									Leu		458
					tac Tyr											506
					gtc Val											554
att Ile 140	caa Gln	ttc Phe	atg Met	gaa Glu	ctt Leu 145	cgg Arg	cga Arg	atg Met	gag Glu	cat His 150	ttg Leu	gtg Val	aaa Lys	ttt Phe	tgg Trp 155	602
					ttt Phe											650
					atg Met											698
					gaa Glu											746
					gat Asp											794
					ctg Leu 225											842
					gaa Glu											890
atg Met	gcc Ala	aga Arg	gca Ala 255	gga Gly	act Thr	cac His	caa Gln	gtt Val 260	tcc Ser	atg Met	gaa Glu	acc Thr	caa Gln 265	gaa Glu	tct Ser	938
	Ser				gta Val											986

						cta Leu 290										1034
aat Asn 300	act Thr	ttt Phe	acc Thr	aaa Lys	tat Tyr 305	ata Ile	tct Ser	cca Pro	gat Asp	gct Ala 310	gct Ala	aaa Lys	cca Pro	ata Ile	cca Pro 315	1082
						aat Asn										1130
						aac Asn										1178
						gag Glu										1226
						att Ile 370										1274
						tgt Cys										1322
						gca Ala										1370
						cag Gln										1418
						gat Asp										1466
						cct Pro 450										1514
						tgc Cys										1562
						cag Gln										1610
						agc Ser										1658
	Ile					gga Gly										1706

	ccg Pro	act Thr 525	gct Ala	cct Pro	ggc	tct Ser	gtt Val 530	ggc Gly	cct Pro	cct Pro	gat Asp	gag Glu 535	tct Ser	cac His	cca Pro	gly aaa	1754
	agt Ser 540	Ser	gac	agc Ser	tct Ser	gcg Ala 545	tct Ser	cag Gln	tcc Ser	agt Ser	gtg Val 550	aaa Lys	aaa Lys	gcc Ala	agt Ser	att Ile 555	1802
	aaa Lys	ata Ile	ctg Leu	aaa Lys	aat Asn 560	ttt Phe	gat Asp	gaa Glu	gcg Ala	ata Ile 565	att Ile	gtg Val	gat Asp	gcg Ala	gca Ala 570	agt Ser	1850
•			cca Pro														1898
	ttt Phe	gga Gly	aga Arg 590	gtg Val	agt Ser	gac Asp	ttg Leu	999 Gly 595	caa Gln	ttc Phe	atc Ile	cgg Arg	gaa Glu 600	tct Ser	gag Glu	cct Pro	1946
•			gat Asp														1994
			tgg Trp														2042
	tgg Trp	aag Lys	att Ile	gct Ala	aaa Lys 640	atg Met	ata Ile	gtc Val	agt Ser	gac Asp 645	gtt Val	atg Met	cag Gln	cag Gln	gct Ala 650	cag Gln	2090
	tat Tyr	gat Asp	caa Gln	ccg Pro 655	tta Leu	gag Glu	aaa Lys	tct Ser	aca Thr 660	aag Lys	tta Leu	tga *	ctca	aaaa	ctt	·	2136
	gaat	29999	gée a	atga icaat	aaaa	a go	acta	tatt	tct:	gato	etgt eagt	tace	gttg	gtt t	ccag	attott gggaga aattgg	2196 2256 2316 2363
	<211 <212	0 > 34 L > 66 2 > PF B > Ho	52	apie	en								٠				
)> 34 Arg	Gly	Ala	Gly	Pro	Ser	Pro	Arg	Gln	Ser	Pro	Arq	Thr	Leu	Arq	
	1		Pro		5					10					15		
	Lys	Glu	Gln	20 Glu	Lys	Thr	Ser		25 Val	Lys	Ser	Ile	Lys	30 Ala	Ser	Ile	
	Ser	Val	35 His	Ser	Pro	Gln	Lys 55	40 Ser	Thr	Lys	Asn	His	45 Ala	Leu	Leu	Glu	
	Ala 65		Gly	Pro	Ser	His 70		Ala	Ile	Asn	Ala 75		Ser	Ala	Asn	Met 80	
	Asp	Ser	Phe	Ser	Ser 85	Ser	Arg	Thr	Ala	Thr 90		Lys	Lys	Gln	Pro 95		
	His	Met	Glu	Ala	Ala	His	Phe	Gly	Asp		Gly	Arg	Ser	Cys		Asp	_

15/122

Tyr Gln Thr Gln Glu Thr Lys Ser Ser Leu Ser Lys Thr Leu Glu Gln Val Leu His Asp Thr Ile Val Leu Pro Tyr Phe Ile Gln Phe Met Glu Leu Arg Arg Met Glu His Leu Val Lys Phe Trp Leu Glu Ala Glu Ser Phe His Ser Thr Thr Trp Ser Arg Ile Arg Ala His Ser Leu Asn Thr Met Lys Gln Ser Ser Leu Ala Glu Pro Val Ser Pro Ser Lys Lys His Glu Thr Thr Ala Ser Phe Leu Thr Asp Ser Leu Asp Lys Arg Leu Glu Asp Ser Gly Ser Ala Gln Leu Phe Met Thr His Ser Glu Gly Ile Asp Leu Asn Asn Arg Thr Asn Ser Thr Gln Asn His Leu Leu Leu Ser Gln Glu Cys Asp Ser Ala His Ser Leu Arg Leu Glu Met Ala Arg Ala Gly Thr His Gln Val Ser Met Glu Thr Gln Glu Ser Ser Ser Thr Leu Thr Val Ala Ser Arg Asn Ser Pro Ala Ser Pro Leu Lys Glu Leu Ser Gly 2.85 Lys Leu Met Lys Ser Ile Glu Gln Asp Ala Val Asn Thr Phe Thr Lys Tyr Ile Ser Pro Asp Ala Ala Lys Pro Ile Pro Ile Thr Glu Ala Met Arg Asn Asp Ile Ile Ala Arg Ile Cys Gly Glu Asp Gly Gln Val Asp Pro Asn Cys Phe Val Leu Ala Gln Ser Ile Val Phe Ser Ala Met Glu Gln Glu His Phe Ser Glu Phe Leu Arg Ser His His Phe Cys Lys Tyr Gln Ile Glu Val Leu Thr Ser Gly Thr Val Tyr Leu Ala Asp Ile Leu Phe Cys Glu Ser Ala Leu Phe Tyr Phe Ser Glu Tyr Met Glu Lys Glu Asp Ala Val Asn Ile Leu Gln Phe Trp Leu Ala Ala Asp Asn Phe Gln Ser Gln Leu Ala Ala Lys Lys Gly Gln Tyr Asp Gly Gln Glu Ala Gln Asn Asp Ala Met Ile Leu Tyr Asp Lys Tyr Phe Ser Leu Gln Ala Thr His Pro Leu Gly Phe Asp Asp Val Val Arg Leu Glu Ile Glu Ser Asn Ile Cys Arg Glu Gly Gly Pro Leu Pro Asn Cys Phe Thr Thr Pro Leu Arg Gln Ala Trp Thr Thr Met Glu Lys Val Phe Leu Pro Gly Phe Leu Ser Ser Asn Leu Tyr Tyr Lys Tyr Leu Asn Asp Leu Ile His Ser Val Arg Gly Asp Glu Phe Leu Gly Gly Asp Val Ser Pro Thr Ala Pro Gly Ser Val Gly Pro Pro Asp Glu Ser His Pro Gly Ser Ser Asp Ser Ser Ala Ser Gln Ser Ser Val Lys Lys Ala Ser Ile Lys Ile Leu Lys Asn Phe Asp Glu Ala Ile Ile Val Asp Ala Ala Ser Leu Asp Pro Glu Ser 565 570 575 Leu Tyr Gln Arg Thr Tyr Ala Gly Lys Met Thr Phe Gly Arg Val Ser

```
585
             580
Asp Leu Gly Gln Phe Ile Arg Glu Ser Glu Pro Glu Pro Asp Val Arg
                                                    605
                              600
        595
    Ser Lys Gly Ser Met Phe Ser Gln Ala Met Lys Lys Trp Val Gln
                          615
                                                620
Gly Asn Thr Asp Glu Ala Gln Glu Glu Leu Ala Trp Lys Ile Ala Lys
                                          635
                     630
                                                                 640
625
Met Ile Val Ser Asp Val Met Gln Gln Ala Gln Tyr Asp Gln Pro Leu
                 645
                                       650
Glu Lys Ser Thr Lys Leu
<210> 35
<211> 162025
<212> DNA
<213> Homo Sapien
<300>
<308> GenBank AC005730
<309> 1998-10-22.
<400> 35
gaatteetat tteaaaagaa acaaatggge caagtatggt ggeteatace tgtaateeca
                                                                             60
gcactttggg aggccgaggt gagtgggtca cttgaggtca ggagttccag gccagtctgg ccaacatggt gaaacactgt ctctactaaa aatacaaaaa ttagccgggc gtggtggcgg
                                                                            120
                                                                            180
                                                                            240
gcacctgtaa tcccagctac tcaggaggct gaggcaggag aattgcttga acctgggaga
tggaggttgc agtgagccga gatcgcgcca ctgctctcca gcctgggtgg cagagtgaga
                                                                            300
ctctgtctca aaaagaaaca aagaaataaa tgaaacaatt ttgttcacat atatttcaca
                                                                            360
aattīgaaat gttaāaggta ttātggtcac tgatateetg ttīcattett tatataatea
                                                                            420
ttaagtttga aatgtatact tgcactacta acacagtagt taatcttagt cctacaagtt
                                                                            480
                                                                            540
actgctttta cacaatatat tttcgtaata tgtatgcact ggtgtttatg tacgtgttta
tgtttatatc tgttaaaatt agcagtttcc atctttttct attttgtacc atcacatcag
                                                                            600
ticagaagga tigacagagc aaaaigattt gatgaagtat aaaagtcaca tggtgagtgg
                                                                            660
cataaataca actotgaaca attaggaggo toactattga otggaactaa actgcaagco
                                                                            720
agaaagacac atateetata tgteaagaga tgtaceace aggeagttaa agaagggaag
                                                                            780
tacacataga aagcacaatg gtgaataatt aaaaaattgg aatttatcag acactggatt
                                                                            840
catttgctcc taaagtcaga gtcctctatt gtttttttgt ttttgtgggt ttcttttaa
                                                                            900
attitttat titttgtaga gtoggagtot cactgtgtta coogggotgg totagaacto
                                                                            960
                                                                           1020
ctggcctcaa acaaacctcc tgcctcagct tcccaaagca ttgggattac agacatgagc
cactgagece ageceagacg ctttageatt tatgaagett etgaaatagt tgtagaaace
                                                                           1080
gcataagett tecatgteae ttteaaagtt tgatggtete tttagtaaae caaccaagtt
                                                                           1140
attcctcaag ggcaaaataa catttctcag tgcaaaactg atgcacttca ttaccaaaag
                                                                           1200
gaaaagacca caactataga ggcgtcattg aaagctgcac tcttcagagg ccaaaaaaaa
                                                                           1260
aggtacaaac acatactaat ggaacattet ttagaagage eccaaagtta atgataaaca ttttcatcaa agagaaaaga gaacaaggtg ttagcaaatt eetetatcaa ataacactaa
                                                                           1320
                                                                           1380
acatcaagga acatcaatgg catgccatgt ggaagaggaa gtgctagctc atgtacaaac
                                                                           1440
cagtagataa tttcaactig ctgccgaatg aaacctcttt gcaaggtatg aatcagcact
                                                                           1500
teleatgitt gittigetti gittigitti gittitagag acaggecett getetgicae
                                                                           1560
acaggetgga gtgcagtgge acgatcagag etcactgcaa ectgaaacte etgggetcaa gggatcetee tgeettagee teccaagtag etgggactae aggeecacca tgeecageta
                                                                           1620
                                                                           1680
attitttaaa titteetatag agatgggate teactageae etticatgit tgatgiteat
                                                                           1740
                                                                           1800
atacaacgac caaggtacaa tgtggaaaag ggtctcaggg atctaaagtg aaggaggacc
agaaagaaaa ggggttgcta catagagtag aagaagttgc acttcatgcc agtctacaac
                                                                           1860
actgctgttt tcctcagagc agagttgatg atctaaatca ggggtcccca acccccagtt
                                                                           1920
catageetgt taggaacegg gecacacage aggaggtgag caataggeaa gegageatta
                                                                           1980
ccacctgggc ttcacctccc gtcagatcag tgatgtcatt agattctcat aggaccatga
                                                                           2040
accetatigt gaactgagea tgeaagggat gtaggtttte egetetttat gagactetaa
                                                                           2100
tgccggaaga tctgtcactg tcttccatca ccctgagatg ggaacatcta gttgcaggaa
                                                                           2160
aacaacctca gggctcccat tgattctata ttacagtgag ttgtatcatt atttcattct
                                                                           2220
atattacaat gtaataataa tagaaataaa ggcacaatag gccaggcgtg gtggctcaca
                                                                           2280
```

		0020000220	acadocadat	caccacctca	ggagatcgag	2340
cotgtaatee	Cagcaccccg	ggaggccaag	gcaggcggac	cacgaggtca	ggagaccgag	2400
accatcctgg	ctaaaacggt	gaaaccccgt	ctactaaaaa	ttcaaaaaaa	aattageegg	
gtgtggtggt	gggcacctgt	agtcccagct	actcgagagg	ctgaggcagg	agaatggtgt	2460
gaacctggga	ggcagagctt	gaggtaagcc	gagatcacgc	cactgcactc	caqcctqqqc	2520
dacadadeda	tactctgtct	Caaaaaaaaa	aaaaaaaaa	aaagaaataa	agtgaacaat	2580
gacagagcga	taccergee	2550000000	2254444444	accondition	cccasasatt	2640
				ccccagttca		
				ccgctaatct		2700
atcttcattc	aatgctaaaa	aatgaataaa	cttttttta	aatacacggt	ctcactttgt	2760
				tagcctcaat		2820
				tacaggcacg		2880
				accatgttgc		2940
ctcagccaat	ccccaaaccc	cotacagaga	coasstacta	cactcaactc	atcctccttc	3000
CtCaaaccct	gggctcaagt	gattettet	caaactcccg	gactcaagtg	*********	3060
cttggcctcc	caaagtgctg	ggattacaag	catgagecae	tgtacccagc	tggataaaca	
ttttaagtcg	cactacagtc	atggacaatc	aggcttttca	acatgcagta	tggacagtga	3120
gtcccagggt	ctgcttttcc	atactgaaat	acatgtgata	ctaaggagaa	aggtgctcgc	3180
aaggatattt	aaaatgaaga	atatttaaaa	tgaggaaaaa	actgtttctt	catgactttq	3240
ataaggetga	taaagaccat	ttctgtgatc	tcaggtgatt	cactcaagta	gtatatttca	3300
ataataatta	tetegascae	cctcaatctt	22C2222t2	ccatgatttt	traatoctot	3360
graarracra	teeggaacag	ccigaaccic	aaccaaaaca	ccacgatete	cadacgetge	3420
tatgatacct	tgatgatatg	accaaactgc	aatgtaggta	gctaaatctc	cacgagetty	
acttccccga	gagttgacag	ttttcttcac	aaattaaaga	aatatattt	ttgatacatg	3480
attggcatat	ttaaaaacta	cactgaaatg	ctgcaaaatg	atataaagaa	acattttcca	3540
gaatcaaatg	caatcaaaga	gtggattagg	aatctactca	ccattatcaa	ctaaatagaa	3600
				gcactttggg		3660
				gcaacatagc		3720
						3780
				gatgcttgta		
				atcaagactg		3840
tggtcatgct	gcgccacagc	ctgagtgaca	gagagagacc	ctgtctcaaa	aacaaaaca	3900
aacaaaaaac	acttaacctt	cctgtttttt	gctgttgttg	ttgttgtttg	tttgttttga	3960
gatggagtct	cactctqttq	cccaggetgg	agtgcagtgg	cgtgatcttg	gctcactgca	4020
				ctcccgagta		4080
t > 0000000000	0035535000	caactacttt	tttacatttt	tagtagagat	gagatttcac	4140
Laggegeeeg		tastasta	cccgcacccc	cagcagagac	ggggcccac	4200
cgtgttagcc	aggatggtct	Lgateteetg	accicgigat	ccacctgcct	eggeeteeta	4260
aagtgctggg	attacaggca	tgagccaccg	cacceggeea	acctttctgt	tttttagttt	
gatatgcttg	ttaactcagc	agctgaaaga	atgctgaaag	tggccttcag	taaaaaaatt	4320
tcactágaat	ctctacatcc	atatttaatc	tgaatgcata	tccagattga	tcagttagag	4380
caaaaacact	catcatcatt	cctgatgacc	tctaattctg	gtttcggctt	tctatttcaa	4440
				tttgtcctgg		4500
				taaaacaagt		4560
ttttaattat	+++++	taagacccct	ttatttcaac	ataaaacata	acttctaccc	4620
						4680
				atcttatcct		
cacagcaatg	caaacatcta	taaaatgcct	ctgcataatt	gtcagaagct	atagtccaga	4740
aatcattgaa	aatgcttttc	cattttaagc	ttaggtgagg	tgtcttagga	aacctctatg	4800
acaacttact	ctatttattg	ggaggtaaac	tcccagactc	tcccagggtc	tcctgtattg	4860
atctcatttt	ttaggcttcc	taatcccttg	aagcacaatc	gaaaaagccc	togatetett	4920
ttctgcacat	atcatcgcgg	aattcattcg	gettecagea	agctgacact	ccatgataca	4980
accected	cccttctcca	aacccccatc	cttactacaa	ttagctagga	taaaaaattt	5040
ageggeeteg		gacgccagtc	cccgccgc	ccagccagga	-503933555	5100
getgggette	agtgcaggct	rergegggtt	cccaageege	accaggtggc	CCCacagget	
ggatgtcacc	attgcacact	gageteetgg	caggctgtac	caatttttta	attatttaat	5160
atttatttt	aaaattatgg	tgaatatttt	ggtattctgc	tctaaaatag	gcccataaat	5220
gcacagcaga	tatctcttgg	aacccacagc	tttccactgg	aagaactaag	tatttttctt	5280
ttaaagatgc	tactaagtct	ctgaaaagtc	cagatectet	acctctttcc	atcccaaact	5340
aagacttgga	atttatgaga	gatchagcha	acagaaatcc	cagacacatc	attggttctt	5400
casacastac	actectecta	andageta	accetaacca	ggcccctgca	ccadaddat	5460
t-t-	ageceeeea	aagaggccca	gccccaagca	ggcccccgca	cta333335	5520
				cagagaggca		
				aacggagggg		5580
actaacataa	agagaaaagg	gaacacagag	aaatggatgg	caccaacaac	cagcaaagcc	5640
				tcatccccaa		5700
				gatcatgtat		5760
				cctatctgct		5820
						5880
ccccagacac	congettada	Latticatia	gurgeragag	acttaaaatg	uacayaacyc	2000

acattgtcag	gatgactatt	accaaaaaat	caaaagacag	caagtattgg	tgaggatgta	5940
				aatggagcag		6000
				ggaaacaact		6060
cactocatoa	aggggtagac	aatatgtggt	atatacatac	catggagtac	tattcagcct	6120
ctasasasas	aaaaaaaaat	tctataacat	gcaacagcat	ggatgaatct	tgaggagatt	6180
						6240
ttgctaatga	aacaaggcag	ccacagaaag	acaaatacty	cacgactcca	ctcatatgag	
				tggaggttac		6300
agggcgggaa	acgaggagtt	actaatcaac	gaacataacg	ttgcagttaa	gtaagatgaa	6360
				caacaataat		6420
cttaaaaatt	tgttaagggt	agattaacaa	atgtagtaga	tccacaaatg	tggttaagtg	6480
ttcttaccac	agtaaaataa	aaaaagaata	tcaagcccag	gagttcgaga	ctagcctggg	6540
taacatggtg	aaaccctgtc	tctacagaaa	atacaaaaat	tagccagctg	tggaggtgca	6600
ctcctaggga	ggctgaggtg	ggaggcttgc	ttgagcccag	gaggtcaagg	ctgcagtgag	6660
				aagacaccac		6720
				gcaagaacaa		6780
				ggtctaactg		6840
				tacacaggta		6900
				gagaaagcgt		6960
						7020
				atcacagaaa		7080
				gaacaccctg		7140
				gctttgcctg		
				caatcaaaga		7200
tcaaattcac	attttagaat	gattttttat	ggcagtatgg	ggaatagatt	aaaagagagt	7260
gaagctggag	gcaagaaact	tgttaagagg	caactgaaac	agtctagatg	ataaataata	7320
				aacagatacc		7380
taacaggact	tgatcaccag	ttgtatcttg	gagaggaagg	agttgtttcc	ttgctttccc	7440
				atatactggt		7500
				tctcccagtt		7560
				agtttctcaa		7620
				tggggccgga		7680
				tgttggcagg		7740
				cctccccatc		7800
cttctctcac	catgageteet	ctalageace	geeggaacae	gaagccacag	tattattat	7860
						7920
				cacttctgcc		7980
				gtgtcaaaga		8040
				tcaattaatg		8100
				acccatagat		
				tgaaaggtaa		8160
				aactattgag		8220
				accatataag		8280
				aacataaatg		8340
gttgacaaag	gtcttggcag	aatcccttac	ccattgactt	ggggccaaga	gagggacact	8400
tctttgtttg	agggataagg	aaaataagaa	agaatgggtg	ctatttagtg	tggtcctgtc	8460
tctagggcaa	acgcataggt	aacaaactgt	gtgtgttagg	aatatagatg	tgacctcaca	8520
				ccttcctacc		8580
				aggtttcagc		8640
				ccttccttgg		8700
				ctgagaacct		8760
				atgtcactgc		8820
				ataaggcgtt		8880
				aaacagtggt		8940
				atgtatgacc		9000
				ttccttttct		9060
						9120
				acaattcatg		
				gtgacatttg		9180
				ggcacatgat		9240
				aaagtgatga		9300
				attcttctag		9360
ttgaagttgt	cacatttaat	ccttaataac	ccactatgag	tcaggtattc	ttctctcccc	9420
tttggacagt	tggggaaatg	ggggtcagag	aggttaggta	atttgctcag	ggccacacaa	9480

```
cctgcatgta gaaaatctga gatttgtaca ggaacgtatc aaactctgaa gtccatgctt
                                                                                     9540
                                                                                     9600
ctattttccc atgctgcctt tctaataaaa ggtaactaat gctactggat gctgcccca
aagtgagtca ctttcacccc accctacttg attttctcca taaaactaat cacatcctga
                                                                                     9660
caacttattt attgctgatc tcccccacta gattataaac tcaataaaag caagatcctt
                                                                                     9720
gtotgotgaa tatoagtaco taaaacgotg totagoacag agcaagtaat taatatttgt
                                                                                     9780
                                                                                     9840
tqaatqaaca aataaaggaa aaaaattcaa aggaagaaaa agccctaaaa cagatgttta
cctaaacata cattttaaaa gaaagcatat aacaaattca ggacagaatt taaatttgat
                                                                                     9900
tttttaaaga aataaccaag tgctagctgg gcacagtggc tcacacctgt aatcctagca ctctgggagg ccgaggcagg cagatcactt gaggtcaaga gttcaagacc agcctggcca
                                                                                     9960
                                                                                    10020
                                                                                    10080
acatggtgaa acctgtctct actaaaaata cagaaattat ccaggcatgg tggcaggtcc
ctgtaaccc agctactcag gaggetgagt caggagaatt gcttgaaccc aggaggcaga ggttgcagtg ggccaagatt gcaccactgc actccagcct gagtaacaaa gcaagactet
                                                                                    10140
                                                                                    10200
10260
                                                                                    10320
                                                                                    10380
aagaaagaaa aagaaagaaa gaaagaaaga accaagtgct tatttgggac ctactatgct
atqtttttcc atqcacqcta ttttcagtaa agcagttagc aaacttgcaa gatcataaca
                                                                                    10440
acaaatatat gcttctataa ctctaaaatt gtgctttaag aagttcctct ttaccagctc
                                                                                    10500
atgtatgcat tagttttcta agagttacta gtaacttttt ccctggagaa tatccacagc cagtttattt aaccaaagga ggatgcttac taacatgaag ttatcaaatg tgagcctaag
                                                                                    10560
                                                                                    10620
ttgggccagt tcatgttaat atactccaga acaaaaacca tcctactgtc ctctgacaat
                                                                                    10680
tttacctgaa aattcatttt ccacattacc aaggagccag ggtaggagaa tatagaaaga ccacccaaga atccttactt ctttcagcaa aatcaattca aagtaggtaa ctaaacacat
                                                                                    10740
                                                                                    10800
gecetaacaa tgaatageag attgtgetea gaagaatgat etacaacate ttaetgtgaa ggaactactg aaatatteea ataagaette tetecaaaat gattttattg aatttgeatt
                                                                                    10860
                                                                                    10920
                                                                                    10980
ttaaaaaata ttttaagcct aaattttaaa aggtttgata ttggtacatg aatagacaaa
                                                                                    11040
cagacatgga ctagaccaag aattaggttc aaacatatac aggaatttaa tatacgataa
atctagtatt ccaaaggaac caacaaatgg tgttcagaca gcaggatagg catcaggaaa
                                                                                    11100
aacacagttg ggcaccctac cttactccta acaccaggag taactgaagg agcaccaaat atttatttat tttaattata gttttaagtt ctagggtacg tgtgcacaac atgcaggttt
                                                                                    11160
                                                                                    11220
                                                                                    11280
attacatagg tatacatgtg ccatgttggt gaggagcacc aaatatttaa aagaaaaaaa
ttggccaggg gcggtggctc acacctgtaa tcccagcact ttgggaggcc aaggtgggca
                                                                                    11340
gatcacctga ggtcgggagt tcgagaccag cctgagcaac atggagaaac cccatctcta
                                                                                    11400
ctaaaaatac aaaattagcc aggcatggtg gcacatgcct gtaatcccag ctacttggga ggctgaggca ggagaatagc tttaatctgg gaggcacagg ttgcggtgag ctgagatatt
                                                                                    11460
                                                                                    11520
                                                                                    11580
gcactccage etgggcaaca agagcaaaac ttcaactcaa aaaaattaat aaataaataa
                                                                                    11640
aaataaaqaa aqaaaaqaaa aaaatqaaaa tagtataatt agcagaagaa aacaccgtag
11700
agaaaatcac ctacatacaa accaaatctt tctacatgcc taaaacatag cacaaacaca gctaaataat catagctgaa tgaactgga aaacaaaact tgactcatat ccagacagag
                                                                                    11760
                                                                                    11820
ttaattttcc tacacataaa gagtacctat ataaacccaa caaaaaaacc accactaacc
                                                                                    11880
caaaataaaa atgtgacagg taatgaacag gtagttcaca gagaatacaa atggctcttc ggcacataag atgctcagac tgacttttac ttatttattt tttgagagac agggtctcac
                                                                                    11940
                                                                                    12000
gatgttgccc aggttaggct caaactcctg ggctcaaatg atagtaccag gactacaggt gtgccccacc gcacctggct cctcaaccac ctgtattaac aggaaatgca aaataaaact
                                                                                    12060
                                                                                    12120
                                                                                    12180
ttcaaatcta ttttacctat tagaatggca aaaatttgaa aaacttcaaa catcatcatg
                                                                                    12240
ttggtgagaa tgtgaggaga ctggcactct cattttttgc tgatagcata tatatactga
tggcttctat ggaaagcaat ctggcagcgt ctatcaaatg tacaagtgca tatatccttt
                                                                                    12300
gacaaagcaa ttecacteta ggaatgtgtt etatatggtt gtgetteetg gggetgggaa etgggageta agggacaggg geagaagata atettettt eceteettee eegttaaaca
                                                                                    12360
                                                                                    12420
                                                                                    12480
tgitgaattt tatatactgt aatatattat ttttcacaaa agataatttt taagcgatat
gtctgggaat ttttttttt cttttctgag acagggtctc actctgtcat ccaggctgga atgccatggt atgatctcag ctgactgcag cctcgacctc ctgggttcaa gcaatcctcc
                                                                                    12540
                                                                                    12600
                                                                                    12660
cacctcagcc tectgagtag etgggactac aggcacgtge catcatgeta attittgtat
                                                                                    12720
atacagggtc tcactatgtt gcccaggcta atgtcaaact cctaggctca agcaatccac
                                                                                    12780
ccacctcagg ctccaaagtg ctgggattac aggcgtgagc caccgcgcct ggccctggga
attettacaa aagaaaaaat atetaetete eeettetatt aaagteaaaa cagagaagga
                                                                                    12840
aattcaacct ataatgaaag tagagaaggg cctcaaccct gagcaacaaa cacaaaggct
                                                                                    12900
                                                                                    12960
atttctgaga caggaatttg ctgaacaaaa tcgagggaag atgacaagaa tcaagactca
cttctcggct gggcgcagtg gctcacacct gtaatcccag cactttggga ggccgaggcg gacagatcac gaggtcagga gattgagacc atactggcta acacagtgaa acccagtctc
                                                                                    13020
                                                                                    13080
```

tootoooot	acaaaaaatt	adccoddcat	aataacaaat	gcctgtagtc	ccagctactt	13140
Caccadadac	~~~~~~~	tacastassa	2252250225	andetteene	taaaccaaca	13200
gggaagctga	ggcaggagaa	Lggcgcgaac	ccaggaageg	gagcttgcag	Lyayeeyaya	
tcacgccact	gcactccagc	ctgggtgaca	gagcaagact	ctgtctcaaa	aaaaaaaaa	13260
aagactcatt	tctctagatc	ttgagccgta	ttcaaattta	tctcagctta	gtgagaggtt	13320
aaagcaagga	atatccttcc	ctgtgggccc	tgctccttac	tgaaggaagg	taacggatga	13380
gtcaaggaca	ccaatqqaga	aaagcactaa	caccattatc	tgatgaacat	tacgtgaaga	13440
agggtaagaa	gtgaagtgga	attoctoaao	aagtcagtga	aagcggacat	tcatttqqqq	13500
aaatooaata	taggaaatcc	ataaaagtga	ttaaaaagat	gttagaggct	aaaacaaaaa	13560
anacagaaca	tcaggaaaatc	gagaggatgg	taactaacac	ggtgaaaccc	catctctact	13620
gaccacaggg	ccaggagacc	agaccaccc	acacacac	ggtgaaaccc	ctactccacc	13680
aaaaatacaa	adaattagee	aggegeggeg	gcaggcaccc	gtagtcccaa	ccaccoggga	
gactgaggca	ggagaatggc	atgaacetgg	gagacggagc	ttgcagtgag	ccgagaccac	13740
gccactgcac	tccagcctgg	gtgacagagt	gagactccat	ctcaaaaaaa	aaagttagat	13800
acgagagata	aagatccaac	agacacacaa	ctgctaattc	tgaacagaac	aaaacaaatg	13860
gcacaggaaa	agaaaattta	agatataaca	ccggaaaact	ttcctgaaat	tgagtaactg	13920
aatctatagc	ttgaaagggt	ttagcatatg	ccaagaaaaa	tcagtagagt	ccaaccagca	13980
caagacacat	ctaqcaaqqc	tggtgattct	accaacacaq	agaaagaagt	gggtgaccca	14040
taatqcqqaa	aaaqqcaqac	catctgcagt	cttctccaga	acactggagt	ctgaagacaa	14100
aagaatgctg	cctactgage	cagaagggag	agaaagtgac	ccaacacatc	tttaccaaqt	14160
tagaatgtca	cocattattt	aaaggctgca	aaagccatga	aagacatgaa	agaacacaag	14220
catttacaac	atgaagaac	acaagcatto	tcatactcaa	gaatccttaa	gaaaaatgta	14280
cacctacaac	acgaaagaac	aaattaaata	tacttaatot	gctcattaat	gadaacttca	14340
						14400
				gcccggcttt		14460
				tgctctcagc		
				tggtctacaa		14520
				ctcctgcata		14580
cctggcctcc	tcactgcccc	tcctgtacat	caggaaggcg	actccttgag	tcttggctct	14640
ggccgcctcc	tccacctgca	gtgagttaac	tcccttacct	actctaggtc	attgctcaaa	14700
tatcaacatc	tcaatqqqqc	cctccctgac	taccctattt	aaattctaca	tactcccctt	14760
gaccccatgg	acctcactca	ccctattcca	cttttattct	tacaatttag	cacttgttct	14820
cttctaacgt	attctaagac	ttactcattt	attacattqt	ttgccacccc	ctctagtaca	14880
taaactccag	aggggaggg	atttctgtct	atttattcat	ttctttatcc	ctaggacata	14940
dascadddca	tagttcagag	tattcaatgt	tatcaatgaa	tgaactagca	gtagtaccag	15000
ttccagttag	gcacacaatt	aaatctaaat	agaattaaat	ctcatggtct	googttaacta	15060
tectagetag	attacatata	attttaagaa	gcctacaaac	aaaaaattaa	taatotaaaa	15120
tygatagaaa	tttagacaca	atticaagaa	statasasas	cactatacat	caaatctcca	15180
acadeactaa	ctigataata	acaacaaaa	ccctgccagg	cactgtggct	tanagratas	15240
ateccageta	ctcaggaggc	tgaggtggaa	ggattacttg	agaccagagt	ttaagattta	
				aaacttaaat		15300
				ggaagtgtga		15360
ctatgctaat	ggttttttt	ttttttagaa	aggtataacc	aaaagatttc	tttctcaagt	15420
cgataaactg	agaaagataa	gcatatcttc	caattaacag	agggggagga	aaagccagat	15480
acaacaaaat	aagatataaa	ttagtttcca	gttgaaaaca	agagtaggag	ttattttgca	15540
				tgataaacag		15600
				aaccacaaac		15660
				acgtgctgag		15720
				agggtcacac		15780
ccectates	tacccactaa	ttacaccac	catatotoac	agagaaactc	agggcactta	15840
gcaccigtaa	caccaccya	atataatata	cacacgegag	tecterece	tocasattas	15900
gagagtataa	caagaacctt	atgucaturg	agacgaggaa	tcctcagccc	coctacted	15960
				acttttgttt		16020
				cacacacaaa		
				catatgtctc		16080
				aagagtgata		16140
cagatggaca	taaaaagctg	tgtgttacgt	gaaaacaaac	tcaagcagca	gcaggatggg	16200
cttatgatag	tcagtatgag	ctaatttctg	gaaaaaaaa	tctagtgtgt	gcacagaaaa	16260
catctgaaag	aacagaaaca	aaactatcag	cagaatattg	agatgtttta	ctaagttgta	16320
				ctttttggaa		16380
				atggtgatgg		16440
				tgtgcgtgag		16500
				taggttctat		16560
				tgcctaaacc		16620
						16680
Laatittata	guccattat	caccacagec	ccaaagccag	actcagcctg	adactacec	10000

21/122

16740 ttcatttgga accettatta aaatgeeaca tacageteet teaaataaaa acaaaceeta ggacctgaca ctaggettee tttgttgeta ctcataatgg ccaagttetg tgettataat 16800 16860 acatettett teattttatt getacatate caagggtttt atatgttttt ettattatat 16920 cttaattcaa aacaccatca cgctcttttc cagatgaaaa taaggaaaag aaattgagca actgactgac ttaaaggtca taaaactata tagtagcaga gtcagcaaaa gaagaaacac 16980 acatetecea agragaget gaaaaccagt accatteace tecagggtga getatataca gattacaaag teacettete taaatgttea aactgaatee catacecata etttaceaet 17040 17100 acctegtaag aacageetca gatetigtta tageettttt titageatge tgaageeaat 17160 17220 aaaatgcttc ccattcagca agagaaacaa gttctgaaac actgaataat ctgcccaggg cctatgaaca tttccactgt gagaaatgtt ctccactgtg tggagaagat ccttactctt 17280 ctccacacag gcagaacatt agaaaaattc ttggattcta tgatgcacag cttaggagtc 17340 tgtttagcac aatttaagtc caaatagtta ttaaatcctc ctctgttcca gaaacagtgc 17400 taaatactgt gaatataaaa attgaaaaga tactctcctg gctcccaaga aagtcagcca 17460 17520 gatagaggag acacaggcac acaaatcact gtcacatgaa getetacete eetaaettea aacgagggcc taagtcacca agaatacagt agcagttgtg actacgagta actactataa 17580 ttcaatactt tatcttccct tagaaaactc ttctcccttg gaaatttatt tgcatttcta 17640 aataccattc cttactaaaa ggaagcaggg ctccttgggg aaatagctga ttctaggtgt ggactatgaa atgaaaatgg tgagtctggg acatcccatg ttgcccagaa atcaaggaac 17700 17760 tgcccaaaga ttaacagagt catgttaaat ggacctaaga gtgaaccaga aggagctcac 17820 titgccccgc gtggaacaat ttcaagaaaa acatgacagt aatgaattat aaaacatgaa 17880 ttaaaataca tattggtact aaaaagagaa caaaaggatg tggctttgga taaagctctt 17940 cttcatggaa gaataccagc taataaatgt aaaggaaatg agagaattag aaaaattatc 18000 18060 attttgtaaa ccttaatata ttcacctaga catgctaaaa ccactgagta aaaggctgct tgggaagagg atgctcacat gatctcagag tttcacacca cagataattt attagataca 18120 ggaaggaaga tgtgatcaag cttcctgtga cccccagcca ggccccacaa cactatgtgc 18180 cteettgtga tgtgggaget acacagcate geccacacag ettetegeca aaactgtttg 18240 aagetaatca caagggaaga actggacage ttetgaceat gagaegetee accagacaac 18300 18360 ttgcttggcc tctccaaaga aacttgcttg gcctctccaa agaaaactca gtttcattta aaaacaaaac taattattta aaaacaaacg aaaagcaagt tgtggacttg agctccaggg 18420 acagagcaga catacttttc cctgttcttc ccagtaagtg gtaataaaaa ccctcaacac tagatataaa acaaatataa gaaggttctg gaaggggaag aggaggcaga ctatccaggt 18480 18540 gccttgagge ceacagaaca acceagtgat gggtteactg ggtettettt ttgetteatt ateteagaet tggagetgaa geageaggea actteaaaae accaagggge acagattgaa 18600 18660 aagccccaag aaaagcctgc cctctctagc caaaggacca ggaaggagac agtctaatga 18720 18780 gatggaacac atttagacag taactgccca tttaccagca ataactgagc agggagccta gacttccagt cttgtgagga cgtaccaagg tacccaacac ccccaccaag gctgagtaag 18840 gactgegact tttatecetg catggeagta gtaaggagee cateceteae cegecageag tgteagggga acetggactt ceaeteceae ceaggagtga tgaggeeete cetgetgggg 18900 18960 tcatgtcaga ggaggcctag tggagattca gtgacttaac cttttcccag agataatgag 19020 gccacctttc ctcctcttc ccccatggtg acagtgaaag cactgtggca agcagtaggc actcctaccc ctcctagcca gggaggtatc agggaggcca agtagggaac cagaataccc 19080 19140 acaaccaccc agcagcaaca ggggtccccc acccattgg gtgtcaatgg aagcagagcg gaaagcctgg atatttaccc ccatctagaa gtaacaagct gatgtccccc ttcttctact 19200 19260 acaatggtgt tcaaaacagg tttaaataag gtctagagtc tgataacgta atacccaaat 19320 cgttgaagtt ttcattgagg atcatttata ccaagagtca ggaagatccc aaactgaaag 19380 agagaaaaga caattgacag acactagcac taagagagca cagatattag aactacctga 19440 19500 aaggatgtta aagcacatat cataagcctc aacaggctgg gcgcggtggc tcacgcctgt aaccccagca ctttgggagg ccgaggcagg tggatcacaa gatcaggaga tcgagaccat 19560 cctggctaac acggtgaaac cccgtctcta ctaaaaatac aaaaaaaat agcaaggcat 19620 ggtggtgggc acctgtagtc ccagctactc gggagcctga ggcaggagaa tggcatgaac ctgggaagag gagcagtgag ccgagatcgc accaccgcac tccagcctgg gcaacagagc 19680 19740 aagacttegt cecaaaaaaa aaaaaaaaaa aaaaaaaage etcaacaaac aactacaaac 19800 gtgcttgaaa caaatgaaaa aaaaatcttg gcaaagaaat aaaagatata tattttggcc aggtgcagtg gctcacagcc tgtaatccct gcactttggg aggctgaggc aggcggatca 19860 19920 cctgaggtca ggagtttgag accagcctga ccaacatgga gaaaccccgt ctctactaaa aatacaaaat tagccagtca tggtggcaca tgcctgtaat cctagctact caggaggccg 19980 20040 20100 aggcaggaga atcgcttgaa ctcaggaggt ggaggttgcg gtgagccgag atcccgccat tgcacattgc actccagect gggcaacaag agcaaaactc catetcaaaa aaatagatac 20160 atattttaat ggaaatttta qaattgaaaa atacagtaac caaattgaat ggaaagacaa 20220 catagaatgg agggggcaga caaaataatc agtgaacttc aacagaaaat aatagaaatt 20280

22/122

acccaatatg aagaacagaa agaaaataga ctggccaaaa aataaagaag aaaaaagagg 20340 20400 20460 atcagggtcc aggaaagaga caaagatggc acagctggaa acgtattcaa aaaataatag 20520 ctgaaaactt cccaaatttg gcaagagaca taaacctata gattcgaaat gctgaaccc aaataaaaag cccaataaaa tccacaccaa aatacatcat agtcaaactt ctgaaaagac 20580 20640 gaaaagagaa aacgtettga aagcagtgag tgaaacaaca etteatgtat aagggaaaaa 20700 20760 caattcaagt aacagattic ttacagaaat taaggaagcc agaaggaaat gacacaatgg ttttcaagtg ctgaaagaaa agaagtgtca acacaaaatt ctagattcag taaaaatatc 20820 cttcaagaat caatgggaaa tcaagacagt ctcagataaa gcaaaataag agaatatgtt gccagcagat ctcccctaaa ggaatggcaa aaggaagatc atgcaacaga ccaaaaaatg 20880 20940 atgaaagaag gaatccagaa acatcaagaa gaaagaaata acatagtaag caaaaataca 21000 tgtaattaca ataaaatttc tatctcctct taagacttct aaattatatt gatggttgaa 21060 qcaaaaatta taaccctgtc tgaagtgctt ctactaaatg tatgcagaga attataaatg 21120 gggaaagtat aggtttctat acctcattga agtggtaaaa tgacaacact gtgaaaagtt 21180 acatacacac acacacgtaa gtatatataa atatatgtgt gtatatgtgt gtgtatatat 21240 atatatacat ataatgtaat acagcaacca ctaacaacac tatacaaaga gataataacc 21300 aaaaacaatt tagataaatt gaaatggaat totaaaaaaat attoaaatao totacaggaa 21360 gacaagacaa aaagagaaaa aaagaggagg acaaactaaa ttttttaaaa acataaataa 21420 aatggtagac ttaagcccta acttatcaat aattacataa atgtaaatga tctaattata 21480 tcaattaaaa gacagagata gcagagttaa tttaaaaaaca tagctataag aaacctgctt 21540 21600 tgggctgagt gcagtgactc acacttgtaa tcccagcact tcgggaggcc aaggcgggtg 21660 gatcacctga ggtcaggagt tccagaccag cctggacaac atggtaatac cccatctcta ctaaaaataa aaaaaatta gccaggcatg gtggcacacg cctgtagtcc caactactca ggaggctgcg acacaagaac tgcttgaacc cgggcagcag aggtagcagt gggccaagat tgcgccactc cagcctgaac gacagagtga gactccacct cagttgaaaa acaaaaaga 21720 21780 21840 aacctgcttt aaatatacca acatatgttg gttgaaatta aaagaataaa atatatcatg 21900 aaaacattaa tcaaaagaaa ggagtggcta tattaataac ataaaataga cttcagagaa 21960 aagaaaattt caagagacag gaataaaagg atcaagaaaa gatcctgaaa gaaaagcagg 22020 caaatcaatc attetgettg gagattcaac accetetett aacaactgat agaacaacta gacaaaaaaa tcagcatgga gttgagaaga acttaacacc actgaacaac aggatctaat 22080 22140 agacatttac ggaacactct acccaacaat agcaaaataa acattctttt caagtattca 22200 ctgaacatat cettagacce taccetggge cataaaacaa agetcaetag tgattgeega 22260 aggettggat ggacagtgga agagetgeat ggggagggag aaggtgacag ttaaagagtg 22320 taggatttct ttttgggata atgaaaatgt tccaaaattg attgtggtga tgttggcgca 22380 actictacaaa tataaaaaaag gccattgaat tgtacgtttt aagtgggtga aacatatggt atgtggatta tatctaacgc tttttaaaaa cttaacacat ttcaaagaat agaagtcata 22440 22500 22560 caqaqtqtqc tctactqqaa tcaaactaqa aaqaqqtaac tggaggataa cgagaaaaqc ctccaaatac ttgaaaactg gacagcacat ttctaaaatc atccgtgggt caaagatatt 22620 catttctgat attcattttt attgtttaat gtatttttaa aaatttctta agggaaataa 22680 actgactaaa aatgaatatg gctgggtgcg gtggctcacg cctgtgatcc cagcactttg 22740 ggaggccgag gctggtggat cacaagatca ggagttcgag accagcctgg ccaagatggt 22800 22860 gaaaccccgt ctcaactaaa aaactacaaa aagtagccaa gcgcagtggc gggagcctgt ggtcccagct acttgggagg ctgaggtagg agaatcgctt gaacacaggc agcagaggtt gcagtgagcc aagattgtgc cactgcacgc cagcctgggc gacagagact gcctcaaaaa aaaaaaaaa aaaaagaata tcaaaatttg tgggacatag ttaaagcaat gctgagaggg 22920 22980 23040 aaatttataa cactaaatgt ttacattaga aaagagaaaa agtttcaaat caatagtctc 23100 cactcccatc tcaagaacac agaagatgaa gagcaaaata aacccaaaagc aagcaaaaga aagaaaatat aaaaataaat cagtaaaatt gaaaacagaa acacaataaa gaaaatcagt 23160 23220 23280 gaaacaaagt actgattett egaaagatta ataaaattga caaaceteta geaaggetaa caaacaaaaa agaaagaaga cacggattac cagttattag aatgaaagca taattagaaa 23340 caactctaca cattataaat ttgacaatgt agatgaaatg gactaattac tgaaaaaaca 23400 caaattacca caactcaccc aatatgaaat agataattgg gatagcctga taactactga 23460 gaaaattgaa tttgtaattt taacactctt aaaacagaaa cattaaactt aatattttat 23520 23580 aaatattaga taaggtaatt ataccettee ttaacaaata aaaacgacaa attattttge agctaaagag atgiatgtac tgtgaaaaat atcttcagaa aaatagaact ttgtttgaag 23640 aataaggatt taaaaaatgt tittaactct caagaagcaa atatcigggc ccagaiggtt 23700 tcactgaaga attctaccaa atgtttaatg aagaattacc accaactcta catagcatct 23760 23820 ttgagaaaac tgaagagaag ggaacatctc ccagttcatt ttatgaagtg ggtgttactc 23880 tgatactaga actgtataag gacagctact cttgacacac tgcctatggg tagctctgct

ctgcaggaac	agtcagaaaa	aaaaaaaaa	gaagcactgg	acaagggcag	tataaaaaaa	23940
gaaaactggg	ccaggtgcag	tggctcacac	ctgtaatctc	agcactttgg	gaggctgacg	24000
ctaataatc	acctgaggtc	aggagtttga	gactagectg	gccaacatgg	taaaaccctq	24060
tototactaa	aatacaaaaa	ttagccaggc	aggatagtag	ggaaaataaa	aaggaaaaaa	24120
22202222	aaactgcaga	ccaatatcct	tcatgagtat	agacacaaaa	ctccttaaac	24180
taattaaaac	aaaccgcaga	actacaccc	atatataaaa	ataattatac	accatoatca	24240
teettaacaa	aatattagta	agragaagca	atacacaaa	ataattatac	accacgatca	24300
agtgggactt	accocagaaa	egeaagtetg	gettaatatt	tgaaaacaag	gcaacccact	
				atcaatgcag		24360
ttgccaaaat	ccaatatcca	ttcatgatac	tctaataaga	aaaataagaa	taaaggggaa	24420
attccttgac	ttgataaagc	ttacaaaaga	ctacaaaagc	ttacagctaa	cctatactta	24480
atggtgaaaa	actaaatgct	ttcccctacg	atcaggaaca	aagcaaggat	gttcactctc	24540
attgctctta	tttaacatag	ccctgaagtt	ctaacttgtg	caaaacgata	agaaagggaa	24600
atgaaagacc	tgcagattgg	caaagaagaa	ataaaactgt	tcctgtttgc	agatgacatg	24660
attqtctcat	agaaaatgta	aagcaactag	gggtaggggg	gcagtggaga	cacgctggtc	24720
aaaggatacc	aaatttcagt	taggaggagt	aagttcaaga	tacctattgc	acaacatggt	24780
aactatactt	aatatattgt	attettgaaa	atactaaaaq	agtgggtgtt	aagcgttctc	24840
accacaaaaa	tgataactat	gtgaagtaat	gcatacgtta	attagcacaa	cotatattac	24900
tccacaaaaa	catottotac	atgataaata	cacacaattt	tatctgtcag	tttaaaaaca	24960
tootttta	acasaasas	atgacaaata	cctctaatcc	cagcatttta	agagggggg	25020
catgattttg	gccaggcaca	grygereata	cccgcaaccc	cagcacccca	ggaggetgag	25080
				ggtcaacata		25140
				gtgcacctgt		
acttgggagg	ctgaggcacg	agaattgctt	gaacaaggga	ggcagaggtt	gcagtgagct	25200
				catctcaaaa		25260
aaagcatgac	ttttcttaaa	tgcaaagcag	ccaagcgcag	tggctcatgc	ctgtaatccc	25320
accactttgg	gaggccgagg	caggcagatc	acaaggtcag	gagtttgaga	ccagcctgac	25380
caacatggtg	aaaccccatc	tctactaaaa	aatatataaa	ttagccaggc	atgtgtagtc	25440
tcagctactc	aggaggctga	ggcaggagaa	tcacttgaac	ccggaggcag	aggttgcagt	25500
attaaaccac	cgcactccag	cctqqqtqaq	agaacgagac	tccgtctcaa	aaaaaaaag	25560
caaaataacc	taattttaaa	aacactaaaa	ctactaaqtq	aattcagtaa	gtctttagga	25620
ttcaggatat	atgatgaaca	tacaaaaatc	aattgagctg	gacaaaggag	gattgtttta	25680
catcagteat	ttgaggctgt	aatgcacaat	gattgtgcct	gtgaatagct	actatactec	25740
ggccagcagc	ccataatcac	accecatate	tatttaaaaa	aaaaaaaatt	gtatctctat	25800
agecegagea	taacaacgag	actacatte	ttaaaaaa	aataaatact	gtttttaatt	25860
gtactagtaa	Laagcacatg	ggcaccaaaa	ctaaaaacac	aacaaacacc	gtttttaatt	25920
gcctgaaaaa	aatgaaatac	ttacatataa	acctaacaaa	atgtgcagga		25980
gaaaactaca	aaacgctgat	aaaagaaacc	aaagaagact	taaatagcgt	gaaatatacc	
atgcttatag	gttggaaaac	ttaatatagt	aaagatgcca	attttatcca	aattattaca	26040
				acatagatat		26100
atacaaaaat	gtatacggaa	atatgcaaag	gaactagagt	agctaaaaca	aatttgaaaa	26160
agaaaaataa	agtgggaaga	atcagtctat	ccagtttcaa	gacttacata	gctacagtaa	26220
tcaagactgt	gatattgaca	gagggacagc	tatagatcaa	tgcaaccaaa	tagagaacta	26280
agaaagaagc	acacacaaat	atgcccaaat	gatttctgac	aaaggtgtta	aaacacttca	26340
acqqqqqaaq	atatgtctct	cattaaaggg	tgtagagtca	ttgcacatct	ataggcaaaa	26400
agatgaacct	gaacctcaca	ccctacagaa	aaattaactc	aaaatgactc	aaggactaaa	26460
				gcacggtggc		26520
aatcccagca	ctttgggagg	ccaaggcagg	togatcacct	aaggtcagga	gtttgagacc	26580
agecogatea	acatogagaa	gccccatctc	tactaaaaat	acaaaattag	ctagacatag	26640
				catgagaatc		26700
cacacacac	acttacata	agecaettgg	gaggeegagg	actoragest	gcccgaaccc	26760
ggggggcaga	ggccgcggcg	agccaagacc	acaccaccgc	actccagcct	ttgggtaatata	26820
				tagaaaatct		26880
	aagaattctt			atctataagg		
	ctgtaattcc			gaaaggtata		26940
				tttagtaaaa		27000
aagacctgac	gcagtggctc	acacctgtaa	tcccagcact	ttgggaggcc	ggggcaggtg	27060
gatcacacgg	tcaggagttc	gagaccagcc	tggccaacat	ggtgaaaccc	cgtctctacc	27120
aaaaatataa	aaaattagcc	gggtgtcatg	gtgcatgcct	gtaatcccag	ctactgagga	27180
				ttgcagtgag		27240
				tctcaaaaga		27300
				gccagaattt		27360
				agagcaacca		27420
actocactto	aacagctgtt	aaatttgcta	actotttaco	aagagagccc	tatcaatatc	27480
~3~3349~~	~~~~			~~3~5~5~		

```
actgtcattt gaggctgaca ataagcacac ccaaagctgt acctccttga ggagcaacat
                                                                             27540
aaggggttta accotgttag ggtgttaatg gtttggatat ggtttgtttg gccccaccga gtctcatgtt gaaatttgtt ccccagtact ggaggtgggg ccttattgga aggtgtctga
                                                                             27600
                                                                             27660
                                                                             27720
gtcatggggg tggcatatcc ctcctgaatg gtttggtgcc attcttgcag gaatgagtga
gttettacte ttagttecca caacaactgg ttattaaaaa cageetggea ettteeccea
                                                                             27780
tetetegett ceteteteae catgtgatet caetggttee cettecett atgeaatgag
                                                                             27840
                                                                             27900
tggaagcage ctgaagecet egecagaage agatagtgat gecatgette ttgtacagee
tacaaaacca tgagcccaat aaaccttttt tctttataaa ttatccagcc tcaggtattc
                                                                             27960
ctttatagca agacaaatga accaagacag ggggaaatca acttcattaa aataatctat
                                                                            28020
gcagtcacta aacaaataag aacaagaggc tccagaagtg ggaagccaat acccagagtt cctacaatac agtatctgaa aagtccagtt tccaaccaaa aaatatatat atacaggccg
                                                                             28080
                                                                             28140
                                                                             28200
gacatggtag cttatgtctg taatcccagc actttgggat gctgaggcgg gcagatcacc
ctaggicagg agitogagac cagootggoo aatatggoaa aaccoogtot otactaaaaa
                                                                             28260
tacaaaaatt agccaggcat ggtggtggat gcctgtaatc ccagctactc gggaggctga
                                                                             28320
ggcagggaat cacttgaacc caggaggcag aggttgcagt gagccgagat cacgccactg aactccagcc tgggcaacaa agtgagactc cacctcaaaa aaaaaaaaa tatacatata
                                                                             28380
                                                                             28440
                                                                             28500
tatatgtgtg tgtgtgtgtg tgcgcgcgtg tgtgtatata cacatacaca tatatacata
tatacagaca cacatatata tatgaagcat gaaaagaaac aaggaagtat gaaccatact
                                                                            28560
ttctgtggtt atgataggat ggggtatcac gggggaagta gacaagggaa actgcaagtg
                                                                             28620
28680
                                                                             28740
                                                                             28800
tccaaataga gaatatcaat aaaggcatag aaattataaa atataataca atggaaattc
cqqaqttgaa aggtagaata actaaaattt aaaattcact agagaaggtt caacactata
                                                                             28860
tttgaactgg cagaagaaaa atttagtgag acaaatatac ttcaatagac attattcaaa
                                                                             28920
tgaaaaataa aaagaaaaaa gaatgaagaa aaataaacag aatctcagca aaatgtggca
                                                                             28980
                                                                             29040
caccattaat cacattaaca tatgcatact gagagtaccg gaagcagatg agaaagagga
                                                                             29100
agaaaaaata ttcaaatgat ggccagtaac ttcctagatt tttgttttaa agcaataacc
tatacaatca agaaactcaa tgaattccaa gtaggataaa tacaaaaaga accacaaaca
                                                                             29160
qatacaccat ggtaaaaatg ctgtaagtca aaaacagaga aaatattgaa agcagctaga
                                                                             29220
ggaaaactta taagagaacc tcacttacaa aagaacatca cttataaaag aaccacaata
                                                                             29280
                                                                             29340
atagaaacag ttgacctctc atcagaaaca atgaatgata acatatttga agtgctcaaa
                                                                             29400
gaaaaaaaat aaagattcct atatacgaca aagctgtctt tcaaaaatat acatccaaaa
ggattgaaac cagggtettg aagagttatt tgtacateca tgttcatage agcattatte
                                                                             29460
acaatagcca aaaggtagaa gcaacccaag ggtccatcga caaataaata aaatgtggta tatgtáfaca caatggaatt tattcagtat taaaaaggaa tgaaattctg acacatgcta
                                                                             29520
                                                                             29580
caacatggct aaaccttgag aacactatgc taagtgaaat aagccagcca caaaaggaca
                                                                             29640
aataccatat tacttcactt gtatgaaata cctagggtag tcaaattcag agatagaaag
                                                                             29700
taaaacagtg gttgccaagg gctgagggag ggagtaacgt ggagttattg ttgaatgggt acagaatttc agttttgcaa gataaaaaga gttctggaga cagatggtgg tgagggtggt
                                                                             29760
                                                                             29820
acaacaatac aaatatactt tatactactg aacagtatac ttaaaaatga ttaacatggt
                                                                             29880
                                                                             29940
gaaaccccgt ctctactaaa aatacaaaaa aattagctgg gtgtggtggc gggcacctgt
                                                                             30000
aatcccagct acttgggagg ctgaggcagc agaattgctt gaaaccagaa ggcggaggtt
gcagtgaget gagattgege cacegeacte tageetggge aataagagea aaacteegte
                                                                             30060
tcaaaaaata aaaaataaaa aaaatttaaa aatgattaag caggaggcca ggcacggtgg
                                                                             30120
ctcacaccta taatgccagc actttgggag gccgaggcag gcgatcactt gagaccagga gtttgagacc agcctggcca acatggcaaa accetgtctc tgctaaaaat acaaaaatta
                                                                             30180
                                                                             30240
                                                                             30300
gccaggcatg gtggcatata cttataatcc cagctactgg tgagactgag acacgagaat
tgettgaace caggaggeag agattgeagt gagtegagat egegeeactg aattecagee
                                                                             30360
tgggcgacag agcaagattc tgtctcgaaa aaacaaaaac aaaaacaaaa agcaaaacca
                                                                             30420
aaaaataatt aagcaggaaa cgagattgct gctgaggagg agaaagatgt gcaggaccaa
                                                                             30480
ggctcatgag agcacaaaac ttttcaaaaa atgtttaatg attaaaatgg taaattttat
                                                                             30540
atgtatetta ecacaaaaa aaggetggg gggeaggaaa tgaaggtgaa ataaagacat eccagagaaa caaaagtaga gaatttgttg eettagaaga aacaceacag gaagttette aggetgaaaa caagtgaeee cagagggtaa tetgaattet cacagaaaat tgaagcatag
                                                                             30600
                                                                             30660
                                                                             30720
                                                                             30780
cagtaaaggt tattctgtaa ctatgacact aacaatgcat attttttcct ttcttctctg
aaatgattta aaaagcaatt gcataaaata ttatatataa agcctattgt tgaacctata
                                                                             30840
acatatatag aaatatactt gtaatatatt tgcaaataac tgcacaaaag agagttggaa
                                                                             30900
                                                                             30960
caaagctgtt actaggctaa agaaattact acagatagta aagtaatata acagggaact
31020
                                                                             31080
gtaatattaa tagacataat acaaaaatac cacaaaaagg gaagaagaca atagaactac
```

ataggaataa	cattttggta	tctaactaga	attaaattat	aaatatgaag	tatattctgg	31140
taaqttaaqa	cacacatgtt	aaaccctaga	tactaaaaaq	taactcacat	aaatacagta	31200
2222221222	taaaataatt	aaaatgttfg	tattagtttc	ctcagggtac	agtaacaaac	31260
	taaataaatt	220202005	aaatatattt	tctcccagtt	ctggaggcta	31320
Laccacaac	chanceter	attacacac	atactcacta	tassaactat	2003233022	31380
aacacctgca	accaaggiga	gracagggcc	acyclotte	tgaaggctct	aggaaagaac	31440
cctcccttgt	ctcttccage	ttccagtggt	tctcagtaac	cctaagtgct	cerragerra	
tagctatatc	attcctagca	accagaaaga	agaaaataat	aaagattatg	gcaaaaaata	31500
atgaaatcaa	aaggagaaaa	atggaaaaaa	ataaataaaa	ccaaaagcta	gttctttgaa	31560
aagatcaacc	aagttaacaa	accttttaac	tagactgaca	aaaaggaggt	aagactcaaa	31620
ttactagaat	cagaaataaa	agagggaca	ttactaatqa	gggattagaa	aagaatacta	31680
ccaacaaatc	totoccaaca	aattagaaaa	cttagatgaa	atggacaggt	toctaggaca	31740
2025022052	ccasaattta	ctcaacaaca	aadadacaat	ttgaatgagc	tataacaagg	31800
						31860
gaagagactg	aactgacaac	caayaaacca	tttataaaga	aaatcccagg	teresana	31920
ttcactgtga	aattetttea	aacttataaa	tataaattaa	catcagttct	teacaaacte	
ctccaaaaaa	aagaacagat	ctctatttac	aggcgatacg	atctttagaa	aatcctaagg	31980
gaactactaa	gacactatga	taactgataa	acaagttcag	caaggctgca	ggatagaaaa	32040
ccaatataca	aaaatctatt	atatttctat	acacttgcag	tgaacaaccc	aaaaatgaga	32100
ttaagaaaat	aattcaattt	acaataacat	caaaaaqaat	aaaaacactc	aaaaataaat	32160
ttattcaagt	aagtgcaaaa	cttatactct	agaagetaca	aaacactgtt	aaaagaaatt	32220
aaacctttac	ataaatmaaa	aactatccca	tottcatooa	tcaaaagact	tattactooc	32280
aaaggtttac	acadacgada	ataaattcaa	casasteett	atcaaaatcc	cadatdagge	32340
aatgettett	adattgattt	acaaacccaa bobook	caaaaccccc	accadaccc	209019095	32400
rgggggrggc	ggillatget	Lycaaccca	gcaccccggg	aggctgaggc	acgcagacta	
cctgaggtcg	ggagetegag	atcageetga	ccaacatgga	gaaaccctat	CECEECLAAA	32460
				cccagctact		32520
aggcaggaga	atcgcttgaa	cccaggaggc	agaggttgca	gtgagccaag	atcgtgccat	32580
				aaaaaaaaa		32640
ccagatgact	tcactgttga	aattgaaaag	attattctaa	aattcacatg	gaattgcaag	32700
accttgagaa	taqccaaaac	aaacttgaaa	aacacgaaca	aaatatagga	tgactcactt	32760
gccaattgca	aatgttacga	cacagcaaca	gtaatcaaga	ctgtgtggta	ctggcaaaag	32820
acacatacat	acatacatat	caatqqaata	taattqaqaq	tacagaaaca	agcctaaaca	32880
tctatggtaa	atacttttct	atttttttt	tttttttt	cttttttgta	gagatagaat	32940
				aagcaatcct		33000
cctcccaaaa	tactaggata	actoocatoa	gccaccacat	ccagcccaga	tgattttcaa	33060
				gatgatcaga		33120
				ctgcgattat		33180
						33240
caccacatcc	ageceaaatg	attittaaaa	aggicaacaa	gaccattctt	ctcaacaaac	33300
aggrerggga	taatcagata	gtcacatgaa	aaaaaaaatg	aagttggacc	the	
ccatatgcaa	aaattaattc	aaaaatgaat	tgatgactta	aacgtaagag	ttacgactgt	33360
aaaactctta	gaaggaaaca	tacgggtaaa	tcttaaagac	gttaggtttg	acaaagaatt	33420
cttagacatg	acaccaaaag	catgaccaac	taaggtaaaa	tagggtaaat	tgtacctacc	33480
aaaatgaaaa	acctttgtgc	tggaaaggac	accatcaaga	aatggaaagc	caaaatagcc	33540
aaggcaatat	taagcaaaaa	gaacaaagct	ggaggcatca	tactacctga	cttcaaagca	33600
acagtaacca	aaacagcatg	gtactagtag	aaaaacagac	acatagacca	atggaacaga	33660
ataaagaacc	caaaaataaa	tccacatatt	tatagtcaac	tgatttttga	caatgacacc	33720
ccttcaataa	atgatactag	gaaaactgga	tatogatatg	cagaagaata	aaactagacc	33780
cctatctctc	accatataga	aaaatcaact	cagactgaat	taaagacttg	aatgtaagac	33840
ccasactat	assactacto	oteoneeace	taannaaaaa	cgcttcagga	cattootcca	33900
						33960
				caaaaacaaa		34020
				aaacaacaga		34080
aacctgtaga	atgggagaaa	atatttgcaa	actatecate	catcaaggga	ctagtateca	
gaacacacaa	gtgactaaaa	caactcaaca	gcaaaaaagc	aaataatctg	gtttttatat	34140
				caaatgtcac		34200
				ttttaaattg		34260
				gctattctgg		34320
agtataaaat	agccaacaaq	tatgaaaaaa	tgctcaccat	cactaatcat	cagagaaata	34380
aaaatcaaga	ccactatgag	atatcctctc	actccagtta	gaatggctac	tatcaaaaag	34440
acaaaataťa	atggatgctg	gcaaagattt	qqaqaaaqqq	gaactcctat	acactgtggg	34500
tagggatgca	aattggtaat	ggccattata	gaaaataata	ctgaggtttt	tcaaaaaact	34560
				ggtatttatc		34620
				ttgcaacact		34680
					J	

gccaagacag	ggaataaatc	taaatgtgca	tcaacagatg	aatggataaa	gaaaatgtgg	34740
	caatagaata					34800
gcaacatgga	tgaacctgga	ggacattata	tttaatgaaa	taagtaaagc	acaaaaagat	34860
	tgttctcact					34920
	cttgggaaaa					34980
	ctatagaact					35040.
	gaagagattt					35100
	ctaattaccc					35160
	ctcataaata					35220
	acccacataa					35280
	atttataata					35340
	aaatgggtaa					35400
	gacacgaaaa					35460
						35520
	tagaatgtag					35580
	tgttggtaag					35640
acgraaagra	atgcagccac	cccggaaaac	agrerggeag	tooccoaac	Lattadatac	
	tatgacccag					35700
	taaaaacttg					35760
	cccatatgcc					35820
	atattatctg					35880
	aaccttatgc					35940
	cggaaatgac					36000
	gctgggagga					36060
	ggggagctta					36120
	aaggtgtcag					36180
tcccaggcct	ctctccttgg	ctggcaggtg	gccatcttct	ccctgcgtct	tcacatcatc	36240
ttttctctgt	gtgtgcccat	gtccaaattt	tgattggctc	attctgggtc	atggccaatt	36300
gctatgcaca	aagtgaagtc	tacttccaaa	agaagggaag	agggaacact	gactaggcta	36360
aacttatagt	cattttaatg	teegetttte	ctatgagatt	gtgaacacac	agaagtaggg	36420
tttttatcta	cattgtgcaa	agtttaataa	gaaaaataga.	attcaagaga	agcagttcaa	36480
tagcaggaat	ttaatatggg	aactaattac	aaggtttagg	gcaggactaa	aaagccagtt	36540
gggatggtga	gccaacccag	agattagcaa	cagtgggacc	ccatctacct	accacccatg	36600
	gataaaggag					36660
ggctggagct	gggaccaccc	tagagacact	gtgcaaagca	gaaaacaagg	gggaaaaacc	36720
ctgacttctc	ccttcctccc	acctttcaat	ctcccactag	tgcttcctac	tagccatact	36780
tggccagaga	cagtgacaag	gaacactgca	aaatgaagtt	tgtaggaatc	atctccctct	36840
	aatatggaag					36900
	atttatcttt					36960
	aactgaaaac					37020
	aaatttattc					37080
	ctataatcag					37140
cacactatac	aacctgaagg	cacaatgcat	gaaaacgttt	caagaaatga	caagatttga	37200
agtcaaattc	taagtgcttt	tccagaatct	ctcaagacga	ttatatagct	acccatttt	37260
	atggaaactt					37320
	ttctggaatt					37380
	ctgacctcct					37440
ttggtctata	gtttacatct	acgggcttat	actottcttt	ttttcatttt	tttaaaattt	37500
ccaaccccca	gtatccatat	actoctctct	atcagggtta	ttttaacttt	cteadatcec	37560
ctgagatgct	ttccatgttt	tttttttt	ttttctccca	catttqaata	gcataggeg	37620
taccaccatc	aaccttggat	tatttaagga	ttcaccattc	cacceguata	ttttttattc	37680
agagtettte	ttgtcattcc	tactatagea	202032000	atctcacctt	tocaoctata	37740
						37800
	atggaatttg					37860
	ccttcattta					
	cctgctgccc					37920
	tgctccctca					37980
tigigicitg	tccatcacta	taatctcagc	acctagtacc	tagtaggtac	ttaccatgta	38040
ticattagca	aaatgttatg	tataaccttg	caccttaaaa	acaagagaag	gaagacaaaa	38100
ttaagtetta	agactatggt	ttagaacatg	gatcagaaac	tacagtctgc	agcccaaatc	38160
	gaagagacca					38220
ggagcagagc	taagtagttc	caagggaaca	cacggccctg	caaagcctaa	aatatttact	38280

```
38340
ctatagetet teacagaaaa agtttteaga teectegttt agaactettg tteatatgea
atttcactaa accatagttt titgggttig tttggtittt titggcaaaa aggaatgagc
                                                                              38400
cgatccagaa aaggttgaaa agaatgaatc attactgctg aaagaatgtg cacacagtcc
                                                                              38460
gicagtatte tgetgeeatg etgacaceea tecaatagtg teatgagatg cageagetae
                                                                              38520
                                                                              38580
tactgtgttc tcaatgccga gtccacccac tccataacca tgtccaagca atcttgggaa
catcatcacc atgettgttt atcettaagg tattgeetca catacagcag tggetggtca
                                                                              38640
                                                                              38700
taaagtcaaa tgacactagt ggccaggagg tcaagagaat gagtgaggac aggtgggtag
gcagcccagg ccctagcaac agcaggagct cacccctcag tcactctagc caggactgaa
                                                                              38760
atacttttca ccctttcaag agagactagg aatctggatt tttatgtgaa atatcttgat
                                                                              38820
tactaaatgt tgtcaacaga catgtcaaaa ggtaaaacta agtaagttca tggggcagat
                                                                              38880
tgactattca ggttatagaa ttaaggattc ttatccaaca cagataccaa ccaaaaagct
                                                                              38940
gacgtataac atattaggag aaactatgtg cactgtcgaa acatcaacaa ggggctaatg
                                                                              39000
totaaaatag totatattgg attocagitg aaacatgggg aaaggacatg aacaggcaac
                                                                              39060
ttatgtcaat ggaaactcaa aaagataaca agcatatata aaagcattct caaattcagt
                                                                              39120
agtaaacaga cagatgcaaa taaaaagagg gaaactgctg ccgggcacag tggctcacac ctgtaatccc agcactttgg gaggccgagg cgggcggatc atgaagtcag gagatcgaga
                                                                              39180
                                                                              39240
ccatcctggc taacatggtg aaaccccgtc tctactgaaa acacaaaaaa ttagccaggc
                                                                              39300
gtagtggtgg gcaccagtag tcccagctac tcaggaggtt gaggcaggag aatggcatga
                                                                              39360
                                                                              39420
acccaggagg eggagattge agtgageega gaccatgeea etgeacteea geetgggega
39480
                                                                              39540
gtaaataaat aaaaagagag agactgctaa agtctagaaa gttgaatgat gccaagcgca
tgcaaagatc agggccttgg gatggccggg tgcagtggct cacgcctgta atcccaccac
                                                                              39600
tttgggagge caaggeggge ggateatgag gteaagagat caagaceate etggeegaca cagtgaaace eggtetetae taaaagtaca aaaaaatata tatatata tatattatta
                                                                              39660
                                                                              39720
tattatatat atatatatca gagocttggg aatocttgtg tgctgctggg gaaggtagtg gtgcagccac ccttgacagc aatotggcag tacttggtta tattaagtat aggcacacac
                                                                              39780
                                                                              39840
cacgaccagg cagtectact cetgggteta aateccaaag aattetcaca caagtecata
                                                                              39900
aggagacatg tacgaggete atteageatt actgggagtg ggaateaace tgggtgteca
                                                                              39960
tctacaggag acgagatgga caaaatgtgg tggatattaa gaccagaatc accaagtaac
                                                                              40020
agagatgggt ggtgagtgac aatcctaaga tacagaataa aggctagaac atgatgccat
tcatgtaaat taaaaataga tgcacacaaa gcagtatacg cgtgaccctt gaatagcaca
                                                                              40080
                                                                              40140
                                                                              40200
ggtttgaact geetgtgtee acttaeatgt ggattttett eeacttetge taeeeceaag
acagcaagac caacccctct tcttcctcct cccctcagc ctactcaaca tgaagatgac
                                                                              40260
aaggatgaag acttttatga taatccaatt ccaaggaact aatgaaaagt atatittete
                                                                              40320
tteettätga ttttetttat etetagetta eattatteta agaatatggt acataataca
                                                                              40380
                                                                              40440
catcacacge aaaataaatg ttaattgact gtttatatta tgggtaagge ttecaeteaa
cagtaggctg tcagtagtta agttttggga gtcaaaagtt atacacagat tttcaactgt
                                                                              40500
gcaggcaatc agttcccctg accccctcat tgttcacggg tcaactgtat atacacaaaa
                                                                              40560
                                                                              40620
gtattatatg aacctcatta gaatagctgt ctatagggag aagagaatga gagtgggata
                                                                              40680
aaacggaatg aacaaataaa ccaacaaatg cattaacaag caaaacaaca gaggggcttg
catgggccag tgatgataaa gggctaagaa tgagaatata attaattcaa ttcctcacac
                                                                              40740
                                                                              40800
ctgaggtcta aaaccaagga aagggagggc caggcgtgga ggctcacgcc tgtaatccca
gcactttggg aggctgaggc gggcggatca caagattagg agtttgagat cagcctggcc
                                                                              40860
aacacagtga aagcccatct ctacaaaaaa tacaaqaatt acccaggtgt ggtggcacat
                                                                              40920
gcctgtagtt agctactctg gaggctgagg caggagaatc acttgaaccc aggaggcgga ggttgcaggg agccgagatc acaccattgc actccagcct gggtgacaga gtaagactct
                                                                              40980
                                                                              41040
                                                                              41100
gtctcaaaaa aataaaaaaa ataaaaaaac agagaaaggg aggaaactag atccaggctg
actagataca gcctttagag ttagaaaaga tgatttgaca atctaagccc acactcagat
                                                                              41160
tgaalgaaat lgaaaagcci ttcaaactaa aacattlaat tacaccatct gctgcagaca
                                                                              41220
gaactcagac aactcaaaca ggtaatgtca gcgtggtgtt ttatatcacc accctcaaca cagaataaaa atcagctgca tgtgaagcag tgactagaat gaagaaaagg ctgcttctta
                                                                              41280
                                                                              41340
ctteetteta gtggttettt eegaaaacat taataggeac eagetetatg catgteacee
                                                                              41400
tgcagggaga catggggtat ataactatga cttactgttc attcctcaag gaattcccaatcttgtggaa gattatacac aatgaggcaa caaaaactat ccaataaaac cacggaaaag
                                                                              41460
                                                                              41520
aagccagtga caaagaagcc agtgatgaaa ggccctgtga gcagagctga tggccatttg
                                                                              41580
gggaagaaag accaacatgg atgggggtga tcagggtggc tccgtgggaa agctggaaga
                                                                              41640
gaagtggcag atototgago tggatgatgg gocactacoa totgtatatg gotaattaaa gacoatgtgt ggatttttta ttoagotott togtgtoatt cotgotatoa goacagaaco
                                                                              41700
                                                                              41760
caatctcaac tttccagcta tattgagcta aacttctcac ctcatggaat ttgcagataa
                                                                              41820
agttcaaaag gatccttgcc ttttcaaaat aattttgaat ggttgagtag tccctctgtg
                                                                              41880
```

ctctctcact	gacaccctct	caaggetget	gagcacgtgc	catgctatgg	ctttctccaa	41940
	tattatass	tangettean	cttaatacaa	atatattata	tetteagaga	42000
				atgtgttctc		42060
				ttctaggtaa		
				gttaacacag		42120
actgcttttt	aaaataaata	aataaataaa	taaataaata	aataaataaa	taaataaata	42180
aataaatgaa	tgatagggtc	ttctgtattg	gccaggctag	tctcaaattc	ctggcttcaa	42240
gagatectee	caccttggtc	tcccacagtg	ttgggattat	agacatgagc	cattgtgctt	42300
				tggttaatcc		42360
cctgggaact	tagatttcag	aagggttccc	accatccaac	ctggaaagag	ggactcactg	42420
				tcttcaggta		42480
						42540
				caataaaaac		42600
gggtetetag	tgagtttece	tggtagacag	catttcacat	gcgttgtcac	ageteetee	42660
ccggggagtt	aagcacatac	accetgigig	actgcactgg	gagaggatgc	Liggaagett	
gtgcctggct	tcctttggac	ttggccccat	gcacctttcc	ctttgctgat	tgtgctttgt	42720
				atgctgagtc		42780
accaccagat	ctgagcatgg	tcctgggggc	ccccaacaca	gaaataaatt	ataaaagacc	42840
aaggactggg	catggtggcc	catgccggta	atctcagcgc	tttgggaggc	cgaggcagga	42900
ggaccagtta	agcccaaaag	ttcaaagtta	cagtgaccta	tgactgcgcc	aatgcactct	42960
				aactaaacac		43020
				ttttaaaaca		43080
				gtctaggaaa		43140
				gcccaataca		43200
				aacagacttg		43260
						43320
				ggaactgcag		43380
				ccactcagca		
				acaggaaaaa		43440
				attcagtatg		43500
				tactagagca		43560
attattagat	attcttggga	agacctaagg	ggacattata	aagagcaagc	agttggtatg	43620
				tgaagaccag		43680
atgctactat	ttgtgcaaaa	aaggagaaat	ggagaatctg	attcatattt	gcttgtattt	43740
qcatqaaqaa	actttggaag	gtacataagt	aactaacaac	aatggttacc	tacttgtaag	43800
				tgtgtccgga		43860
				gtgagcgtaa		43920
				gtgttcggag		43980
				cagacetteg		44040
				cctggtgagt		44100
				gttgcagctc		44160
						44220
				aacatcaaaa		
tcagcagcgc	ggaatgcgac	cgcagcacgt	taccactctt	ggctcgggca	geetgettt	44280
				cattttacag		44340
				agctgattgg		44400
cagagtgctg	attggtgcgt	ttacaatccc	tgagctagac	acagggtgct	gactggtgta	44460
tttacaatcc	cttagctaga	cataaaggtt	ctcaagtccc	caccagactc	aggagcccag	44520
ctggcttcac	ccagtggatc	cggcatcagt	gccacaggtg	gagctgcctg	ccagtcccgc	44580
				gactgggcgc		44640
				caggaggtgg		44700
				aggcagctaa		44760
				ctcactgcct		44820
				gcccacgccc		44880
				cccgcgcctc		44940
						45000
				agcccagaaa		
				gagtgggcac		45060
gaggcaccga	gagcgagcga	ggactgccag	cacgctgtca	cctctcactt	ccatttatgc	45120
				ctatttttt		45180
				tggtgccatc		45240
gcaagctccg	cctcccgggt	tcacaccatt	ctcctgcctc	aacctcctga	gtagctggga	45300
				tattttttt		45360
				acctcgtgat		45420
				cgccctgcct		45480

tttcaaaagt	taaactttaa	gaagtagaaa	cccgtggcca	ggcgtggtgg	ctcacgcctg	45540
				aggtcaggag		45600
tcctqqttaa	cacagtgaaa	ccccqtcqct	actaaaaata	caaaaaatta	gccgggcgtg	45660
				gcaggagaat		45720
				ccttccagcc		45780
				cccggaaagt		45840
				gttgacggaa		45900
				ccagaatgca		45960
				ggaattgcct		46020
				tgggcaacac		46080
				tatttgagag		46140
						46200
				ttcccaagta		46260
				acttctccag		46320
				tgagcaggaa		46380
				tcatgcctgt		46440
				gtttgagatc		46500
				gctgggtatg		
				cacttgaaac		46560
				ggtgacagag		46620
				attttagaac		46680
				actacaaaga		46740
agacgcgtgc	acttcttcct	agataccagc	agataaagca	atatctccaa	aattcagaag	46800
				atggaaaagt		46860
				cacagaatcc		46920
aaccaaataa	tcacttaagg	actcattaag	aaaacaaatg	aaataaaagc	accaatgatg	46980
agtaaataat	cagaaaaatt	tacagtttac	ctaaataact	gtttatgcat	aatgtatgaa	47040
aacccaaaaa	tttaatatgg	gacagaatta	aaatcatgat.	aagattcttt	tttgctttac	47100
tcatggagag	ttcacataaa	cagattatct	tttaatagca	agagaaaaaa	atgtttagat	47160
atgtgtgaaa	aactaagggt	accaaaacag	tgcaaattca	tttatcatca	ggaaaatcca	47220
aattaaaacc	acagtatcca	ccagaataac	taaaaggtaa	aagacagaaa	ttaccaagag	47280
ttggcaagaa	tgtggagcaa	ccacatatac	ttctggggta	aataagttgg	tgcaaccggt	47340
actgaaaact	gtttgctagt	atctactaaa	accgagcaca	tgcacagact	acaaccaagc	47400
agttccactc	ccagatacac	actcaacaga	aatgcacaca	ctcactcaac	aaaagacgtg	47460
tactagagtg	ttcatgtact	tactattcat	aatagtccaa	aaatgcaaac	aaccaactgc	47520
				aatggcatat		47580
				gtaatctcag		47640
gggtaaggca	ggcagatcac	ttgaggtcag	aaatttgaga	ctagcctggc	caacacggtt	47700
				tagtggttgc		47760
				acccgaaagc		47820
				tagagcaaga		47880
				atccgcaaaa		47940
				caaagatttc		48000
				tagcaaggat		48060
				cagcctctgc		48120
				gatccaacaa		48180
				atatttgtac		48240
				ataaatatcc		48300
				attaattagc		48360
				agggcattac		48420
						48480
				atattagata		48540
				tggtggttgc		48600
				ttctggagat		
				aaagtagtaa		48660 48720
				atatttatta		
				agagtaacaa		48780
				atataaatgc		48840
				aaataagatg		48900
				gaacctttaa		48960
				aacatagaaa		49020
acccctacaa	gaatcataat	gggagtcttc	aatacaactc	tccatatcaa	caggtcaaac	49080
						<i>p</i> .

agagaaaaaa	aataagttaa	ggatgcagaa	aacctgaatt	accatcaata	aacttgagat	49140
				aacagtcgtg		49200
				agcactgtgg		49260
atcctaaaaa	ccttccagct	acaaaacatc	tttttaaaaa	tataaaaaaa	tacaaaaata	49320
accccaaaaa	castacaaca	cacatootoa	aaccaaaatt	ctagaataca	cccacaataa	49380
						49440
				gctgaagtaa		
				tctgtgatac		49500
				aatgttgacc		49560
caaaaacaat	agcaataatg	tctcgttgga	agggttgaag	taaaaataca	attaaggcca	49620
aatgtgaggt	aagtggaatg	aaagaattag	aagtccttgc	cttgttcaca	ggactgatta	49680
aataaatgag	ccaggttttc	cattcaaaca	gttaaaactt	gaacaaaata	aactcaaatt	49740
				ataaaaaatc		49800
atcaataaat	cctggttaat	aaaagctggt	tctttgaaag	gattaataaa	ataatcatta	49860
				aagtactgta		49920
				tacaccagaa		49980
				gattttctag		50040
				aataaatgat		50100
				acaccagcag		50160
						50220
agcotgcagt	caagicity	caaacttgag	ggaacagaca	attettetat	ttettateaa	50280
agaaaatgat	ggaaagttte	ccaatttaat	cagagaggac	agcctgatcc	tegetacgaa	
				accaaaaccc		50340
gctagcttat	tgatgtgaac	aatccaaaag	tgcattttaa	attagcccag	ggttttagag	50400
aaagaaaatc	tagcaatgtg	accaccactt	atgttaacaa	ttttaagacg	aaaatctaca	50460
tgatcatatc	aatgcatgct	acacaaaagc	atttgggcaa	aaaacccaac	acccaccctt	50520
gactttttaa	actcttagta	attaggcata	aacagaaatg	tacttaatgt	gatagaatac	50580
actcggtgaa	gatacagagg	gaatgctccc	taaaaccaag	cccaagacaa	agattcctat	50640
				gaagacaagg		50700
				acctacttgg		50760
				gtatggaggc		50820
				agttaatcct		50880
				cctcctcccc		50940
				tccaaaaaca		51000
taaaaaaaaa	acttotocco	ttactcaaaa	aacaguutet	ctgctgcctg	carctectae	51060
agttggtgg	ggataagget	CCCSSCCSCC	tcaccataat	tcttcctgca	aagaccaaaa	51120
						51180
				ccgcttatta		51240
				agtcaaactt		51300
				catgggcact		
				aacccacggg		51360
				agctgacccg		51420
				gagcgcggcg		51480
				aaggcacaac		51540
				ccgcgcacct		51600
ggcgagggct	ggggtcgcct	ccagggccgc	agctgtcggg	agccacctgg	ctctcagtcc	51660
				cggccacctg		51720
ctgcggcacc	ggtcccaccg	ctccgggccg	ggcaggacag	gccaggacgt	ccctcctggg	51780
ctggggacag	gacacgcgac	gaggggaccg	gggccccgc	ggcgaagacg	cagcacgcct	51840
				gacccccgcg		51900
				cggcgggcgg		51960
				aacccccagc		52020
				gccggcgcct		52080
				agatgtagta		52140
				ctctctccca		52200
				cgatgtttgg		52260
						52320
				ttctaagggt		52320
				gattctccgt		-
ctattcggtt	Liccotgcac	taaaaccatg	aaggtggggc	ccagcagtcc	acattetege	52440
				gctatgctca		52500
teegeeeege	aaggagcccg	gtcttgcctg	tggcgctagc	cgcacacgga	cacctcatcc	52560
tgcggggccc	gccccccgc	tgcaccctca	ccgcccaacg	cctcctccgg	gatgcagcgg	52620
aggegeetgg	aagtcggcaa	ggtcaacatc	cccctcagca	tcttccctac	cctcacggct	52680

31/122

cctcctccag gggtgcctca tggccagggg ttagaaagag ccactgtgtt tcttgacatg gaagtggcct aagaccttaa tgaaaactgc aggagtggaa tgacagaacc tttggtcata 52740 52800 52860 cttgagggcg tgaagctcaa atgaggagga aggaaaggat ccagggagaa taaccaaccc tggcaagttg tggcgcccag gtagagggc gagcctaggc tagcggttct cgaccagggc cggtgttgcc cctcctcgcc gccccgcgta catttgggga ggtctggaga cattttggt 52920 52980 tgtcatgatg cgggagttgc tactgttgcc taagtgggta gacacgaggg tgctcctcaa catcctacct gaaggacagg actgcccac aaggaagaat gatccggccc caaataagaa 53040 53100 accetggget ggtcagcaac aaccetttg ttetgagaag agaggaggaa agaataaaag 53160 aagtggggtg aagttttggt ttggtagagg aaacttgaag acattttcac tggaaaggaa gagaggaaga ggagggagat gtctgtaagg acgagcaaac cgggtgacag ctgatttcct 53220 53280 53340 cataligaag taatgagtoo tagttataat aaattootaa taaaaaccca gtttatoodt gcaataaact tgtctttttt ttttaaatat actgcttgat tctgtttgct aatattttat 53400 ttacaggett tgeattgata tgeaaaaatg agatgggeaa taattttett tttgaatgte 53460 taatgttgtt tggtttcaga atcaatgtta tgctcacatc ataaaaaatt tggaaccgag 53520 gcaggaggag tgcttgaggc cagaagttcg agaccagtct aggaaacaca gtgagacccc 53580 cccatctcta caaaaaaaaa aaaagaaaaa aaaatgggca tgtttgcttt ttccttttac 53640 tctgaacaat ttaaggagca ttaaaattat ctattctttg aggtttgatc atttcccagt 53700 taaaaatgtt cctccagcc tgatgctttc tttggggagg gtaaatcttt taaggctaga aaagtttctt ctgtggcaat tttattattt acattttaaa aattattcta gagttaattt tgataaagca tgtatttctt aaaacaaatt atcctttttt tccagatgtt caagtgtatt 53760 53820 53880 53940 tgcataaagt tgaggaaagt agtottttgt gaatotttta acttotocca aatatottat titgtgtait titgcttott tattttgtta acttttaaaa gtgtattttt ttttcaaaga 54000 atcagetett aggittatgt tittggitat actggagett tittettett ettittaaaa 54060 tattittet cetttatti ttagaegtat titgatetaa egtaategga agaaggtaaa ttagaatett tigitaetat tgigittita titeteetta titetetgaa gieetgetti 54120 54180 54240 ataaatagta ccatgttatt tgtgcataaa tattcatttg tcttatattc ttgggaattt toccactica toataaaatg accttoottg totcatttaa tgtgttcaaa otttgcootg aatttaactt tgtotgatat tttaccatco tgotgaattt tgtttgttac occaaacaac 54300 54360 ctttgctgtt ttcgtctttt ctgaaccctt tattttaggt aatcccttga attagagcac taagttttgc tttgtgatta aatctgaaaa tctttatctt gccatagatg agttgagccc 54420 54480 54540 taticatgtg acagetatat tatgetgttt catagecett ttggteettt tttcactett gcattgcata ttttgtgttt attgtgtttt gtgtttcttc tgataatttg gaaggtttgt 54600 atttttattc agggagttgc cttataatca tactccgcaa tacacatcgt cctcagtttc 54660 54720 ttcagactgt ctgttaactc cctattctga ataaaaatga cattgtaatt tccctctttt ttetttacce etttettet eetcacetaa tgtaaatgat tttateette tttagtattt 54780 54840 gettttttaa ttaactacat ttataaatat etttateact tgatttttaa atcagetttg aatgagatat ttggatteet agatataaaa gatgttaatt ataccattte caegttagta 54900 ggtttataaa atcatacatt ctgctgtgta accataatcc cacgtttgtt ttagttccac 54960 tectacagtt aaaagattea gaagtattat taacagttat tttgecatag tttttteece aaeceatttt gtggtaagtt atgateetge tttagtttet taagaataat ttatagagea 55020 55080 55140 gagtgtggtg gctcacgttt gtaatcccag cactttggga gacaagaggt agaaggatcg cttgaagcca gcagttcaag accaccctga gcaacatagt gagaccttgt ctctacaaaa 55200 aattttaaaa tttagecaga egtagtggeg tgtgeetata gteecageta etcaggagge 55260 tgaggcaaga ggattgctag agcccagaag tttgaggctg cagtgacctc tgattgtgcc actgcacccc agtctgggca agaaagtgag aacctatctc tttaaaataa caataataac 55320 55380 ttatgaaaat tatatteeet gagtttttea tgtttaaaaa tatttgttge etttateetg 55440 taaaagtttg agtataaatt cttgggttat actttattta ttgaagaatg tataagtatt 55500 gtettetaga attgagtgtt getgtaatga aaccagaagt cageetggtt tattttteet 55560 cagaaatgag gtaattgccg gccggacacc gtggctcatg cctgtaatcc caacactttg ggaggccgag acaggtggat cacgaggtca ggagattgag accatcctgg ctaacatggt 55620 55680 gaaacccogg ctctactaaa agtacaaaaa gttagctggg catggtggtg gacgcctgta 55740 atcocageta coegggagge tgaggcagga gaatggegtg aacetgggag gaggagettg cagagagetg agategegee actgeactee ageetgggeg acagagtgag acteegtete 55800 55860 55920 aaaaaaacaa aaaaaaaca aagaagtgaa gtaattgcca tgatgctcca agaattatct 55980 ctttgtctat gaaatccaga aatctcactg ttatacattt tggaattatt attctgggcc aatattteet gggacacaat agattgaete tatagattta atttttttt ttttttgag 56040 acagagtete actgeaatet cagettaetg caacetetge etcaegggtt caageaatte teetgeetea geeteceaag tagetgggae tacaggegeg tggeaecatg eetggetaat 56100 56160 ttttgtcttt ttagtagaga cagggtttca ccatgttggc caggctggtc ttgaacgcct 56220 aacctcaagt gatccacctg cctcagcctc ccaaagtgct gggattacag gcgtgagcca 56280

ccategoccay cotcaattoc tottectad gatgetaautt totgaagtty aaaacatty 56400 coccatett gotattadg tottectaa gatgetaaa goaccetatt cocagetage 56400 coccatett gotatgage toggetygt oftecaaga gototgytt totectygets 56400 coccatett gotatgage toggetygt oftecaaga gototgytt tottectygets 56520 tittittytit tigacayagt cottactory caccaaga gototgytt toggataatt 56520 cagetaact gaacatetyg cacagetyge cagetaatty tigagetygt typacatetyg cacagetyge gotataatty 56640 tagetoggat tacaagage gototgytty gotataatty 56640 tagetoggat tacaagage gototgytyge gotatagaga 56700 coctaageto cocaagatyge gogatataatagagagattit cocagetyge cacagetyge typacatetyge cagetygyge toggaccata tittigtatit tragatagaga 56700 coctaageto gotattyge gagattacag gotytageca cocgooctag categyting 56880 gotatetygt gatatitita aaaaaaaaaa gaagtocaat tittigtatit traatititit 56880 gotatetygg gatatitita gatecocat actaacaa gaagtocatt tittagattit 57000 gotytgeac tottoggagetyge toggatygaget toggagagat toggagagagagagagagagagagagagagagagagagag							
ccccactit getaglagae teggeteget citectaaa gatetgtaaa geecetatt eccagetag 56400 aattetaagt acctetgtea geetecact ggtttatgat ttggagtttt tteteetget 5640 aattetaagt acctetgtea geetecact ggtttatgat ttggagtttt ttggtttttg 56580 cageteactg caacetegt teccaggt ggtaattat 56580 cageteactg caacetegt teccaggt tgagegate tectgeetea geetactatga 56600 taggtggttea ecatgtgge cagggtggt ttgagegate tectgeetea geetactatga 56700 tggggtttea catggtgge ggcacaca eccaggtagt ttttgagtt ttaggagttt ttagtagtag 56900 tgggggttea ecatggtge ggagtataag ggtggagea gacggagtgaggtgatttatagt ggtgagtatatt ttagtttta gattagaggagt ggagagaga	ccatgcccag	cctcaattcc	tctttctatc	: tggtaatttt	tctgaagttg	aaaacatttg	56340
ccccatctt gctagtage tcgstggt cttacacaaga gctctggtt tctcctect 56620 ttttttgttt ttgacagagt cttactctgt caccaaggt gagagagt ggcataatct 56520 ttgaggtttact gaacctctgt ctccaaggt gagagagt ggcataatct 56520 taggggtttact cacaggtgg ggcatcacaca ccggctatt tttggttttt ttagtaggg 56640 taggggtttac cacaggtgg ggggggt ttgagagagagt ggcataatct 565640 taggggtttac cacaggtgg gagagggt ttgaaccacac 56760 cctcagcctc ccaaagtgc gagattacag ggctgagcca ccgcgcctgg catggttgg 56680 taggtttaatc tgtagttta ataaagaaaga ggcttatggt tggtttctt ttagttgg 56680 tctctcagact gtaatttgta agatccccat atttactt tttggttttg 56680 tctctcagac tgtttattt atttattt atttattt atttattt tattgttggt tcaagagagag tctcccaggttc gaagataacacaca gaagtccat ttcaattctt 56800 fctctcagac tgtttattt atttattt atttattt atttattt tattgttgg 56820 fctcccaactt ttaattttt tagagacag tgcgcacaca tgcgcacat gcaacctcg 57000 fctgtgtcac tctctggagg ttgaagagag gctccccag ttaagtggg ttaaggagg ttgggggggggggt tcaagagagag gctcacaca tgcccaagtga ttacaggca 57120 fctgcaacttt ttaattttt tagagacaga fctcggtgg tcaagcacaca catggcaca catagacaca catgccaca tgccaagtga tctcaaggca 57120 fctgcaacttt taatttatt atttaattt tagagataaca gtctcggtg tcaagcacaca catggcaca catgagagaca gtagagagag gctggagtgg 57180 faatcctcct acctcaagcc ccaaaattac caaggatacac gcccaagta atttaagt 57300 tagagtatgtg actaagagaa gagggtgaca gtagagagag gcgagagag gcctggagagagagagagagagagagagagagagagagag	ttctaatacg	ttatttcagt	gttcttctaa	gatgtgtaaa	gcaccctatt	cccaqqtcaq	56400
aatotcaagt acctotgtoa gottocacot ggtttatgåt Étggagtttt ttggttittg 56520 ttttttgttt ttgacagagt ottactotgt accacaggot ggaagagag ggaatagag 56500 taggtgagttaa gaacotctgt otcocaggtt tgagcgatto tototgotca gottactgag 56640 taggtgggtttaa cacatgttggo cagggtggtt ttgaacotca gactcatgg 56600 tggggtttaa catgttggo cagggtggtt ttgaactot gacctcaggt aatocacctg 56760 ctocagcot ccaaagtgot gagattacag ggtgagcac ecgegotag catggtggt 56820 agttttaatc tgtagtttta ataaagatag tgottatgtt bgtgttott atatttettg 56840 tlottcagac tgtttattt atttattt atttattt tatttattt tattttattt 56940 tlottcagac tgtttattt atttattt atttattt tatttattt tatttattt 56940 tlottcagac tgtttattt attttattt attttattt tatttattt	ccccatctt	gctagtgagc	: tcggctggtt	cttcacaaga	gctctggttt	tetectaett	56460
Ettittigitt ittigacagagi citactotg caccaggit ggagagagagi ggataaatci 56580 cagcitacatg caacctotgi citaceagit tagaggatti tectigotta gottatatgag 566700 titagaggatti taatatgag titagaggaggat titagagagagggggit titagactoti gacctagggat aatccaccig 56700 cotoagooti caaagggg gagagtaaag ggatgagat titagaggaggtitta titagagagaggggggtitaatata titagatgagaggggaggaggagagga							56520
cagctcactg caacctctgt ctcccaggtt tgagcgatt tcctgctca gcctcagga \$5640 tgagtgggtttca cccatgttggc cagggtggtc ttgaactcct gacctcaggt aatccactg \$6760 ctgaggtttta ccctatgtggc cagggtggtc ttgaactcct gacctcaggt aatccactg \$6760 ctccagctc ccaaagtgct gagattacag ggtgagcac cegggctgg catggttttg \$6820 agttttaatc tgtagtttta ataaagatag tgcttatgtt tgtgttttt atatttctt \$6880 ttaatcacaca gaagtcatt ttcaattttt \$6880 ttaatcacaca gaagtcatt ttaattttt \$6880 ttaatcacaca gaagtcatt ttcaatttt \$6880 ttaatcacaca gaagtcatt ttcaatttt \$6880 ttaatcacaca gaagtcact gaaactccg \$7000 tctcccggt tcaagaatt ctcctgcct agcctcccag gtagctggg atacaggac \$7000 tctcccggt tcaagaatt ctcctgcct agcctcccag gtagctggg atacaggac \$7000 tctccccggt tcaagaatt ctcctgcct ttgagcagc tgagctggg actcaagaca gtccccacaca tgcccagtta attttaatt \$7000 tttttgtagaag acagggtcc catatgttgc caggttgcct tctctagcct tgggagtggg actcaagaca caggctcc catatgtgc caggttgcc attttaatt \$7000 tttttgtagaag acagggttc catatgttgc caggttgccc acttatagtgac acatggacaca gcctgtactcc agcattgtga attttgagag acatggagaca gtcaccaca tgtaattgat catccaaacc \$7420 tttttgtagaag acagggtacac gagtaacctg acctagcaca \$7420 tttttgtaatt ttctgggaaa gaaggatac gtcaccaca acttatagga gcctggaact agcatggaa gaatcagaagaa aggtcaccaca acttataggaag cactggaacca gcctgaaccaca gcatggtca gactgagaca acttaggagaca gcctgaaccaca gaattggagaca gaattggagaca gaattggagaca gaattggagaca gactgagaagaa gagataactgt agcacacaca acttaagacaca gaattgaacacacacacacacacacacacacacacacaca	ttttttgtt	ttgacagagt	cttactctqt	cacccagget	ggagagcagt	ggcataatct	
tagggtttea catagtigge caggiggsgte tigaactoci gacctaagg aatccactg 56700 cctoagcotc ccaaagtgct gagattacag gottgagca cccgogctgg catggtttg 56700 cctoagcotc ccaaagtgct gagattacag gottgagca cccgogctgg catggtttg 56820 agtittaat tigtagtttta ataattcttt 56820 gottgatat tigtagattetta gagattacat gottgagtgtt that gottgattgt the statttett 56820 gottgagtgcag tigtagtggg tigtagtgggg taattcaga tigtagtgag tigtagtggg tigtagtgggg taattcaga tigtagtgag tigtagtgggg tagtagtgggg tagtagtgggg tigtagtgggg tigtagtgggg tigtagtgggg tigtagtgggg tigtagtgggg tagtagtgggg tigtagtgggg tigtagtgggg tagtagtgggg cottagtagtag catggagtagtagtgag tagtagtggg cottagtagtag catggagtag agtagtgag agtagtgggg cottagtagtag catggagtag agtagtgag gottgagtggg catggagtggg cottagtagt catggagatagtgagagagatagtgagagagatagtgagagagagataggagagagagagagagagagagatagtga							
Egggtttca ccatgttggc cagggtggtc ttgaactcct gacctcaggt aatccactg 56760 cctcagcctc ccaaagtgct gagattacaa gotytagcca ccgcgcctgg catgtttgg 56820 agutttaatc tgtagtttta ataaagatag tgcttatgtt tgtgtttctt atatttctt 56840 ttcttcagac tgttttgg gaattccaat atctaacaaa gaagtccatt ttcaattctt 56940 ttcttcagac tgttttggagg tgagtgcagg tggcggagt tcaggtacat tacatccact 57060 tctcccgggt tcaagcaatt ctcctgccc agcctcccga gtagctggga ttacaggcac 57120 ctgccactt tttaattttt tatgacaga gtctcgcttt gttgaccagg ctggagtggg 57180 gtggtgtgac tcttgagggattact catacactcc agcctcccga gtagctggga ttacaggcac 57120 ctgccactt tttaattttt taggacaga gtctcgcttt gttgaccagg ctggagtggg 57180 gtggtgcaat catggctgac tataacctcc aatccctgg ctcaagtgat cctcctggct caggctctc agcttctggc ctctaagtgat cctcctggct caggctctc agcttctg gttgaccag ctggagtggg 57180 ttttgtagag acagggtctc catatgttgc catgggccac taggacaca tgccacqgta attttaatt 57300 ttttgtagag acagggtctc catatgttgc cagggtcgc ctcaagtgat cctcctgac 57240 ttttgtagag acagggtaca tttcaatgt tagggtacac gtgaggatac gccttggccc gccttgtcct actatgttaa tttcaaggaa gagggagata gcgtcaccc actatatagt gcttgtgctt 57600 tggattgtg aatttggag ttctatttgc aggcttatt cactagggaca gccttgatcc gccttgatcc gcctcgcac gcaatgaac gttcacccc actatatagt gccttgtctt 57600 tggattgtg aatttggag ttctatttgc aggcttatt cacacgggac gccttgatcc gccttgacca gcattgcca gcctcgcca gcaatgac gttcaccac cctaatgaca gcctcacacac 57780 tggattgttg agttcccaca cctgattcacacac gccacacacac 57780 actttgagat tctctcccag gcgtgttcac gcgattta acctcaggac acctccag 57720 tacttagact cctcccacac cctacacacacacacacacacacac							
cctcagctct ccaaagtgct gagattacag gcgtgagcca ccgcgcctgg catggttetg 56880 gtactcttgg gtaatttgta agatcccat atctacacaa gaagtccatt troattctt 56980 gtactcttgg gtaatttgta agatccccat atctacacaa gaagtccatt troattctt 56980 gtactcttgg gtaatttgta agatccccat atctacacaa gaagtccatt troattctt 56990 gctgtgtcac tctctggagcg tggagtgcag tggcggatc tcaggtgcag tggagtgcg 57000 gctgtgtcac tctcagcgc tggagtgcag tggagtgcag tggcggatc tcaggtgcag tggagtgcag tctcaggcac 57000 gctgtgtcac tctaattttt taggacaag gtctcgcttt gttgaccagg ctaggagtgcg 57000 gctgtgtcac tctaattttt taggacaag gtctcgcttt gttgaccagg ctaggagtgcg 57180 gtgggtgaat catggctgac taggacaca gtctcgcttt gttgaccagg ctggagtgcg 57180 gtgggtggaat catggctgac tataaccacagca catgccaca tgcccagtta attttaatt 57300 ttttgtaga gacagggtgac catactggag catacaggcac catacagcaca tgcccagtcaca taggatactct acatacttaa tttcatatty taggatgacac tgtaattgat catccaaaca ggaatactt acatacttaa tttcatatyt taggatgacac tgtaattgat catccaaacac 57480 gcgttgttt acatacttaggg tcatacttgagt gattggagt gaggtgagacacct acttatagtg tggttggattg gatttggagt tcatattyc cacgggtctc tggagcccc tctgatccc 57660 gccctgccca gcaatgctac cgttcatcca acgggtct caggagtacacct ggagaccac tctgccaca gggtcatct cgttgattg aatttggagt tctatttyc cacgggtct tcggaccct tcagtcacta 57780 gcttcactgc cgtcttccaca gctccacacacacacacacacacacacacacacacacaca							
gateltting ghaattitga apatecceat atetacacaa gaagtecatt treattitting getactiting ghaattitigs apatecceat atetacacaa gaagtecatt treattities for the treatting graattities apatecceat atetacacaa gaagtecatt treattities for the treatting graattities apatecceat atetacacaaa gaagtecatt treattities for the treatting graattities for the treatting graattities for the treatting graattities are graattiled to treat graattities for the treatting graattile graattile treattiles graattiled treattiles for the treattile treattiles graattiled graattiles graattiled graattiles graattiled graattiles graattiled graattiles graattiled gra	cctcagcctc	ccaaagtgct	gagattacag	acatasaccs	cccccagge	categetttee	
gtactcttgg gtaatttgta agatcccat atctacacaa gaagtccatt treaattct tettettttaga tigtstagt to 57000 gctgtgtcac tettegagac tggaagtcag tggagtgag tggagtgatt taggtgacat gaagcactcg 57060 gctgtgtcac tettetgagac tggagtgag tggagtgag tggagtgact taggtgacat gaacactcg 57060 gctgagtgacat tettetgagac tggagtgag taggagtgac taggagtagac 57120 etgcacattt ttaattttt tagagacaga gtctcgctt gttgaccagg ctggagtggg 57180 gtggtgaat catggagacat catgagaca catgagagaca gtaggagtac catgagagacat catgagacat catgagacat catgagagacat catgagagacatagagagacatactgagacatctgagacatactgagacatactgagacatactgagacatactgagacatactgagacatactgagacatactgagacatactgagagacatactgagagacatactgagagacatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagatactgagagatacatac	acttttaatc	totactttta	ataaacatac	tasttatat	tatatttatt	catggtttgg	
ctcttcagac tgtttattt attttattt attttattt tatgttgag atgagatcc 57060 ctcccgggt tcaagcaatt ctcggaggc tgagatgaga	gractettag	ataatttata	acadagacag	atetacacac	gagatacat	atattttttg	
Sctegagia teaagcaati ciccigacic agciccocag staactegga tacagcaatic ciccigocic acciccoggi teaagcaatic tectoscicci agciccocag staactegga tegagiagiagi 57120 cigcactit thatitit tagagacaga gictocigit gitgacagg ciggagiagi 57180 gitgigigaat catagiciga cataaccicci aaatectiggi cicaagiga ciccigactic gitgaciagia catacagica cataaccicci aaccicciggi cicaagiga ciccigacia cicigacia catagitgi cataaccicci aaccicciggi cicaagiga catecagica catagitgi cicciacicci gigcaccaci 57240 cagcitici accaciccic catagitic catagitic catagitic catagitic catagitic catagitic catagitic catagitic accaticaacci 57420 gigcactic actacitica atticaatit tagagitata agcatigagi accacicca 57420 aggatatit ticciggaaa gaggitaaca gigcaccaci actitatagit gigcitigica cicciacacci 57420 cigaagititi aatticgagi ticciattic aggiticacci actitatagit gigcitigica cicciacacci cicciacacic accaggititi aatticgagi ticciacicia aggiticiaci cicciacacaci cicciacacic aggiticiacia aggiticiacia cicciccaci cicciacacic cicciacacagi accicigacia cicciacaci gigcititi aatticiacia cicciccaci cicciacacic cicciacacic cicciacacic accidicacaci giacicciacia cicciacacicia cicciacacic	ttcttcacac	tatttattt	agattttatt	atttatataa	tatattatt	ctcaactcct	
ctgccaggt tcaagcaatt ctcctgcctc agcctcccga gtagctggg ttacaggcag 57180 ctgccaacttt ttaatttttt tagagacaag gtctcggttt gttgacaag gtgggtggg 57180 cagctctctt gatatttttt tagagacaag gtctcggttt gttgacaag gtgggtggg 57180 ctagtgtgaat cataggtga catacaggac catgcacaca tcccagtta attttaattt 57300 cttttgtagag acagggtct catagttgc catagtgac ctccaagtta attttaattt 57300 cttttgtagag acagggtct catagttgc caaggtggc ctcctactcc tggcctcaag 57480 caggatacttct actacttaaatt ttcatatgt taggtgaca tggatatgac gtagtagac accatgcaca 57480 aggatactgt aagaatgaaa gaggtgaca gtagtatgat gctgggacta gcattgtgca 57540 ctgagaattat ttctgggaa gaggagaaa gggctaaca gtagtatgat gctgggacta gcattgtgca 57600 tggattgttg aatttggaag ttctatttga aggctact tagatgagaca gccttgatcc 57600 tggattgttg aatttggaag ttctatttga aggcttattt caactgggca gccttgatcc 57600 tggattgttg aatttggaag ttctatttga aggcttattt caactgggca gccttgatcc 57600 tggattgttg agttcccac cctccacac cctcaaaagag taaccaggaa tcctgcccag 57840 aggctcatctgc 20tctccac cctccacac cctcaaaagag taaccaggaa tcctgcccag 57840 aggctcatctgc cgtcttccga ggcttttca accatggtca gccttgatcc 57780 ggctctactgc cgtctcccac cctccacac cctcaaaagag taaccaggaa tcctgcccag 57840 aggctcact agttcccaa ctggagata acctttgag caaccacacac 57840 accttacaag gtactccaa gtagctgcaa gagagataa acctttaaga gcaaccacacacac 57840 acctcacaca gtagctgcca acctggagat cattggagat cattggagat caaccacacacacacacacacacacacacacacacac	actatata	ttctcacccc	tagagtagag	taggagasta	targuitgag	atggagtete	
stggstgcaat catgsgtgac tataacctcc aaatcctggg tccaagtgat ctcsgcty 57240 cagcctcctg agtagctggc ataaacctcc aaatcctggg tccaagtgat ctcsqcct cagcctcctg agtagctggg actacaggca catgcacca tgcccagtta attttaattt 57360 taatcctcct acctcagcct cccaaattac taggattata agcatgagc accatgccca 57420 gccttgttct actactttaa tttcatatgt taggstgacca tgtaattgac accatgccca 57420 aggatactgt aagaatgaaa gaggctgaca gtagtatgat gctggagcta gcattgtgca 57540 ctgaatgttg aattggagat tccaagctga gtagtatgat gctggagcta gcattgtgca 67540 ctgaatgttg aattggagat tccaagctgac aggstaatcag tgcatggtgca 27540 ctgaatgttg aattggagat tctcatttga gagcttattt caactgggca gccttgatcc 57660 gccctgccca gcaatgctac cgttctctcc accgggtctc tgggacccc tcagtcacta 57720 cacttactac gtccccac ccccacact cctaaaaag taaccaggaa tccgccca 57780 ggtctactgc cgtcttccgt gggctgttc agttcctatt acccagagtc aaactcccag 57840 cattcactac cgttcccac cctcccacac cctaaaaag taaccaggaa tctgcctca 57780 cattcactac ttctgtccct actgggtttc agttcctatt acccagagtc aaactcccag 57840 ggtctactga ttctgccca cctcccacac catgagtta acctcttcag gccaactccc 57900 cactttgcat ttctgtccct actgggtttc agttcctatt acccagagtc tcttgcctca 57860 tcaaccaggc tgtggtgcaa tgacgccatc tcggctaa gccaactccc 57900 cactttgcat ttctgcccc acttgggtga ttacaggatt taccaggat tcttgcaggt 5800 tcaaccaggc tgtggtgcaa tgacgccatc tcggctcact gcaactccc gccaggctg 5800 gcccagctaa ttttttat tttat ctttattgt gcaagctgtga ttacagacag gccacaccac 58140 ggcgtgagca actggccca gccaatttag ccaattata gaccaaaaa actccaaa 38260 ggcgtgagc actggccca gccaatttag ccaattata gaccaaaaa actccaaaa 38260 ggcgtgagac actggccac ccttggcaa taggcagac actggcagac 3620 gccaggctaa tttgtgccc ttttgggag ccaagaggg Ccaagacaa ccaacaacaa 4840 atactcactg accccaata aaatccaaa aaatccaaa accaaaa tccaaaaca 38260 tcaagcctat gagacgaaa gagaaaaa acacaaaa tccaaaacaa 4262 tcaagaccag gagacgaaa gagaaaaaa accaaaaa tacaaaaca 4262 tcaagaccag gagacgaaa gagaaaaaa acaacaaa tccaaacaa agacgcggg 5860 tcaagaccag gagacgaaa gagaaaaaa accaaaaacaa accaacaacaa agacgagaacaa 6262 tcgagaccag gagacgaaacaa cccatgagac ccaacaacaa accaacaacaa acaacaacaacaa acaaca	tetececet	teaaggagge	ctestass	cggcgcgacc	ctaggicact	gcaacctccg	
gEggigcaal catggctgac tataacctcc aaatcctggg ctcaagtgat cctcctgct 57240 cagcctctct gataactgga actacaggac catgacacaa tgcccagtta attttaatt 57300 ttttgtagag acagggtct cataattac taggattata acactacca tggcctcaagt 57420 gccttgttct actacttaa tttcattatt taggattata acactagacca catagacca 57420 aggatactgt actactgaaa gaggctgaca gtagtatgat gctggaacca gcattgtgac 57420 ttgagattgtt tctgggaaa gagggtgaca gtagtatgat gctggaacca gcattgtgac 57420 gcctggactatt tctgggaaa gagggtgaca gtagtatgat gctgggaacca gcattgtgac 57540 ttgagattgtt aatttggagt ttctatttgc aggcttattt caactgggac gccttgatcc 57660 ggcattgcca gcattgtac gcttcccac cctcccactc cctaaaagcg taaccaggaa tcctgcacta agttcccac cctcccactc cctaaaaggg taaccaggaa tcctgcacta cttgagtcac ggttctctgc gggtgttc agttcctat acctagaga tcctgcacta ctgattcccac ctgattccac catgaggttcac tggagggat accttagag gcaactccc 57900 cactttgcat ttctgtccc atacttagt catgaggat accttcatag gcaactccc 57900 cactttgcat ttctgtccc atacttagt catgaggat accttcatag gcaactccc 57900 cactttgcat ttctgccct atacttagt catgaggat accttcactag tcttggcca tgagtgat catctaga gcaactccc gcaccaccac ttctgacct gtggtgcaa tgactcaca gtagctgtga ttacagacag cctctgggt 58000 gccaggat tattttttt cttttagag agacaggtt tacacagggt tctgggtgat tctttatta cttttagtag agacaggtt tacacaggat gccaacacac 58140 gccaggtaat cctttttttt ctttttata ctttttata gaccaaaaa accaaaa accaacacaca aaataccaaa aaataccaaa aaataccaaa aacacacac	ctccccgggt	ttaatt	tagagagaga	ageeteeega	gtagetggga	ttacaggcac	
cagectectg agtagetggg actacaggca catgecacca tgeccagta attitatit 57300 thittgtagag acagggtete catatgttge catgetgge etectactee tgecetcaag 57360 atatectet actateaget cecaattae taggattata ageatgage accatgeea 57420 gectigitet actactitaa titleatatgt taggigacca tgiaattgg cattgigge 57480 aggatactgt aagaatgaaa gaggetgaca gtagtatgat getgggacta geattgigea 57540 etgagattat titetgagat teteattge aggetacet actitatagig tgettgiet 57600 gecetgecca geattged eggteacet cetetacet tetagetgetge 57600 gecetgecca geattged eggteacet cetetacet tetagetgetge 57600 gecetgecca geattged eggteacet cetetacet tetagetgetgetgetgetgetgetgetgetgetgetgetgetg	ctgctactt	caacttcc	tagagacaga	greregerer	gttgaccagg	ctggagtgcg	
tattgtagaag acagggtct catatgttgc ccaggttggc ctcctactc tggctcaag 57360 satcctctct acctcaagcct ccaaattat taggtgacca tgtaattgat catccaaacc 57420 gccttgttct actactttaa tttcatatgt taggtgacca tgtaattgat catccaaacc 57480 aggatactgt aagaatgaaa gaggtgaca gtagtatgat gctgggacta gcattgtgca 57540 ctgaagtttt tctgggaaa gcaggagata cggtcacct acttatagtg tgcttgtct 57660 tggaattgttg aatttggaagt ttctatttga aggcttattt caactgggac gcctgaatcc 57760 gccctgccca gcaatgctac cgtcctccc acgggtctc tgggaccct tagtcacca 37720 tacttagctc agttcccac cctcccactc cctaaaagcg taaccaggaa tcctgcacca 57780 ggtctactgc cgtcttccgt ggggtgttt agttcctatt accagagta tcctgcctca 57780 ggtctactgc cgtcttccgt ggggtgttt agttcctatt accagagta tcctgcctca 57780 ggtctactgc cttccgt ggggtgttt agttcctatt accagagta tcctgcccca 57800 cactttgcat ttctgtccc atatctagt catggagta acatttcatg tctttgagtc 57960 tactacaaag gtaaattttg ctgttttta atttttttt tgagaatggag tctttgccct 57900 cactttgcat ttctgtccc atactcagc tccgggaga agcccaccacacca	gragigaar	catggetgae	tataacctcc	aaateetggg	ctcaagtgat	cctcctgcct	
gadatcetect accteaged cecaaattac taggstatata ageatgaged accatgecea 57480 aggatactgt aagaatgaaa gaggetgaca gtagtatgat getgggacta getattgat catecaaacc 57480 aggatactgt tetgggaat teteggaagt cagggagata eggteacct acttatatagt 57540 teggatattat tetegagat tetegatette agtetacted tetggagat tetegagat tetgatette aggetatett gaggetatette gegatetace geaatgetac ceteccacte cetaaaageg accettgatet 57660 geetgeetgeetgeetgeetgetgetgetgetgetgetge	cageeceetg	agragerggg	actacaggca	catgccacca	tgcccagtta	attttaattt	
gocttyctic actactitiaa titicatatyt laagitgacca tigtaattigat catccaaacc 57480 aggatatyt aagaatgaaa gaggctgaca gtagtatyat gotggacta gottytyt 57600 tiggattyt aattiggaagt tictatitig aggettattt caactgggca gottygatcc 57600 gocttycaca gaatgctaa gytictotic actitatagty tiggaccoct 57600 tiggattyt aattiggaagt tictatitig aggettattt caactgggca gottygatcc 57600 goottycoca goaatgctaa gytictotic actitatagti tiggaccoct toagteact 57720 tactiatagti agticoccaa gytictocaa cottocaatic cotaaaagga tiggaccatic toggatcaact 57780 gyticaacca citiggatic agticoctaa actitiggat accitigatic agticoctaa tottocaacaa citiggatic cagagettia accititiggat accitigatic agticoctaa titiggatic agagettia accititiggat accitigatic 57900 tactitiacaaa gaaattitig cititititia attitititi tiggatiggag caacticoc 57900 tacticaagaci tictigocta atactiaga gagetgaaa agacagggt tiggagacaa tititititia citititigata agacagggti ticaacaagga tititititia citititigata agacagggti ticaacaagga gocaacacca 58140 goccagata tititititia citititigaa gaacagggti ticaacaagga gocaacacac 58140 goccagataa tititititia citititigaa gaacagggti ticaacaagga gocaacacac 58140 gocgagatagaa gagataaaa aaatocaaa aaatocaaaa aaatocaaa aa							
aggatactgt aagaatgaaa gaggctgaca gtagtatgat gctggcacta gcattgtgca 575400 tggattgtt tctgggaaa gcaggagaata cggtcaccct acttatagtg tgcttgttct 57660 gccctgccca gcaatgctac cgttctctcc acctgactct tgggacccct tcagtcacta 57720 gccctgccca gcaatgctac cgttctctcc accggactct tgggacccct tcagtcacta 57780 ggtctactgc cgtctcctcc cctccacact cctaaaaggc gaacccct tcagtcacta 57780 ggtctactgc cgtcttccgt gggctgtttc agttcctat acccaggact caccactaggagt caccactaggagt caccactaggagt caccactaggagt caccactaggagt tcactaggagt catcaggagt caccactaggagt gcaaccccc 57900 tactacacaag gtaaattttg ctgttttta atttttttt tgagatggag tcttggcctg 57960 tactacacag gtagatttg catcagcact tcggtgcaa tacggcgcact gcaaccccc 57900 tactacacag gtagatttg catcagagt tagacgagat tcactaggagt tcaccaggagt gcaaccccc 57900 tactacacag gtagatttg catcagaggagt tagacgaggaggaggaggaggaggaggaggaggaggaggag	taatcctcct	acctcagcct	cccaaattac	taggattata	agcatgagcc	accatgccca	
tggattatat tictgggaaa gcaggagata cggtcacct acttatagtg tgcttgtctt 57660 gccctgccca gcaatgctac cgttctccc accgggtctc tgggagcccct tcagtcacta 57720 tacttagctc agttcccac cctcacact cctaaaagcg gccttactgg ggctgtttc agttcctatta acttcacacta ggactcactactg gggctgtttc agttcctatta acccaggat acctggcact atttcctact ctggtctca attcttagt caggagttc attttcctatt acctagagat accactactac gccaactactac cgtactcacactactactactactactactactactactact	gccttgttct	actactttaa	tttcatatgt	taggtgacca	tgtaattgat	catccaaacc	
ggattgattg aatttggagt tectatttg aggettatt teagaged teggactgate 57660 geettgacca gaatgetaa egittetetee aceggatete teggacacet teagtacata 57720 tacttagete agtteceate cetaaaageg taaccaggat tectgectea 57780 aacttegete cgtetteegt gggetgttte agtteetatt accetaggaat acceagage tectgateeg 57840 acceagaget tettgateeg acttggage cagagettta accetateage gecaacteeg 57800 tacttacaaa gaaatttg etgtgtttta attititit teggagegge teggagegeggegggeggggggggggg	aggatactgt	aagaatgaaa	gaggctgaca	gtagtatgat	gctgggacta	gcattgtgca	57540
goctgocca goaatgotac ogtotecac acceggitote teggaaccet teagtcacta 57720 tacttagote agttoccaac octocaacte cotaaaageg taaacaggaa tectoccag 57840 gototecate cigattocag acttegagte acattecate accetation gocaactetacetacetacetacetacetacetacetaceta	ctgagattat	ttctgggaaa	gcaggagata	cggtcaccct	acttatagtg	tgcttgtctt	57600
ggtctactgc cgtcttccg gggtcttcact cctaaaaggg taaccaggaa tcctgctca 57780 ggtctactgc gggttttc agttcatgt acccagagtc aaactccag 57800 cacttcctac ctgattcag acttgagtc cagagcttta acccatgagac aaactccag 57800 cacttttgcat ttctgtccct atatcttagt ccatgagagat acatttcatg tctttgagtc 57960 tacttacaaa gtaaatttt ctgttttta tttttttt ttgagagagt tcttttaggtc 57960 tacttacaaa gtaaatttt ctgttttta tttttttt tgagatggg tcttgcctg 58020 gccaagctac tttttttat ctgttttta tttttttt tgagatggg tcttgcctg 58020 gtcttgaatt cctgacctc gccactcac gtagctgca ttaccagagat tacacagcag caccaccac 58140 gccaagctaa ttttttttat cttttagtag agacagggtt tacacagtgt ttacaggagc gcaccaccac gcagctact ttttttat cttttagtag agacagggtt tacacagtgt tgagattaca 58260 ggcgtgagc actgtgccca gccaatttt gagaattaca tacacaga aaatccaaa aaatccaaga cccacaactaa ccacaacttaa cacaaactta cacaaactaa cacaactta cacaaactaa 58380 tactcaggct aaccccagaa ctttgggaaa cgggaaataa cacaacttaa cacaaaataa 58500 aactcaagtc gagaggaaa tggggaaac cacacacga cctggcaaat aggggaaaa cacacacgg ggagaacacac cccaatggga ggagaacaca cacaacttaa cacaaaataa ggcagaaaca cacaacttaa cacaaataaa ggcagaaaca ttgagagaga cacacacac ggagaacacac cacaacttaa cacaaaataa ggcagaacacac cacaacttaa cacaaacttaa cacaaataaa ggcagaacacac gagacacacac cacaacttaa cacaaaataaa ggcagaacacac gagacacacac cacaacttaa cacaaaataa ggcagaacacac ggagaacacac gagacacacac cacaactacac ggcagaacacac ggagaacacac cacaactacac ggcagaacacac gagacacacac cacaactacac ggaacacacac cacaactacac ggagaacacacac cacaacacac ttcgggaacacacacac cacaactacacacacacacacacacaca	tggattgttg	aatttggagt	ttctatttgc	aggcttattt	caactgggca	gccttgatcc	57660
ggtctactgc cgtcttccg gggtcttcact cctaaaaggg taaccaggaa tcctgctca 57780 ggtctactgc gggttttc agttcatgt acccagagtc aaactccag 57800 cacttcctac ctgattcag acttgagtc cagagcttta acccatgagac aaactccag 57800 cacttttgcat ttctgtccct atatcttagt ccatgagagat acatttcatg tctttgagtc 57960 tacttacaaa gtaaatttt ctgttttta tttttttt ttgagagagt tcttttaggtc 57960 tacttacaaa gtaaatttt ctgttttta tttttttt tgagatggg tcttgcctg 58020 gccaagctac tttttttat ctgttttta tttttttt tgagatggg tcttgcctg 58020 gtcttgaatt cctgacctc gccactcac gtagctgca ttaccagagat tacacagcag caccaccac 58140 gccaagctaa ttttttttat cttttagtag agacagggtt tacacagtgt ttacaggagc gcaccaccac gcagctact ttttttat cttttagtag agacagggtt tacacagtgt tgagattaca 58260 ggcgtgagc actgtgccca gccaatttt gagaattaca tacacaga aaatccaaa aaatccaaga cccacaactaa ccacaacttaa cacaaactta cacaaactaa cacaactta cacaaactaa 58380 tactcaggct aaccccagaa ctttgggaaa cgggaaataa cacaacttaa cacaaaataa 58500 aactcaagtc gagaggaaa tggggaaac cacacacga cctggcaaat aggggaaaa cacacacgg ggagaacacac cccaatggga ggagaacaca cacaacttaa cacaaaataa ggcagaaaca cacaacttaa cacaaataaa ggcagaaaca ttgagagaga cacacacac ggagaacacac cacaacttaa cacaaaataa ggcagaacacac cacaacttaa cacaaacttaa cacaaataaa ggcagaacacac gagacacacac cacaacttaa cacaaaataaa ggcagaacacac gagacacacac cacaacttaa cacaaaataa ggcagaacacac ggagaacacac gagacacacac cacaactacac ggcagaacacac ggagaacacac cacaactacac ggcagaacacac gagacacacac cacaactacac ggaacacacac cacaactacac ggagaacacacac cacaacacac ttcgggaacacacacac cacaactacacacacacacacacacaca	gccctgccca	gcaatgctac	cgttctctcc	accgggtctc	tgggacccct	tcagtcacta	57720
ggictactactg cgicticogi gggctgitti agticotatt acceagagit aaaactoccag 57840 cattocotac ctagatocag actiggaga caagaggaga acattiticatg ccatgagat tottigact 57960 tactocacaggi titotogic gaaattitig ccatgagat acattiticatg tottigagit 58020 tacaccaggi titgggtgaa tgacgacate teggstacat gaagaccace ctcotogggt 58020 tacacgagat tittitat cittitatiat cttitagiag agacagggt titacagacag gccacaccac 58140 gccagataat cattititat cttitagiag agacagggt titacagacag gccacaccac 58140 ggcgtgagcc actggccag gccaattitig cttititat attitititat cttitagiag agacagggt titacagacag gccacaccac 58140 ggcgtgagcc actggccag gccaattitig cttititat attitititit gagatggag gccacaccac 58140 ggcgtgagcca titacagacag gagacagggt titacagacag ggcagagt gagataaca actcggccc ccaaagtgc titacagacag ggcgggtaggcc actggcgtgagcc ccaaagtgc titacagacag titacagaag aaatccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaatacaaaat cacaacat tacaaaati cacaacati cacaagac 58500 tacacgagctgg aacccacaa gagagaaaaa acatcaaaat tacaaaati aaaatacaa agaaataac cacaacati aaaatacaa agaaataac cacaacati aaaatacaa agaacaaaca cacaacati agaagagaaa tagagagaaa tagagagaga tacacaaaggggggggagaataacagggggggagaataacagggggggg	tacttagctc	agttccccac	cctcccactc	cctaaaaqcq	taaccaggaa	tectacetea	57780
cattecetac ctgateccag acttegagate cagagettta acetetteag gecaacteee 57900 tacttacaaaa ttotegecet atatettagt ceatingagat acatteeag tettingagee 57960 tactacaaaaa tagetgagat tettingagee 57960 tactacaaaaa tagetgagat tettingagee 57960 tactacaagae geaacteee geaaceee geaaceee geaaceee 58140 gecaageaa tettititat gaategagat teacaagaea geaceaceae 58140 gecaagetaa tettititat ctititagaa agacagggtt teacaagaea geaceaceae 58260 gegeggagaea gegggagaeae actingeee geaaceaee 58260 gegggagataage eegggagaee atteggagee attegggeee geaaceaee gegggagataage teacageee geaaceaee gegggagaeae acceeeae geaaceaee geaaceaee geaceaeee 58260 gegggagataage eegggagaee atteggagee attegggeee attegggeee deceaageee teacaaaaaa aaateeeaaa aageaacaaa cacaacatea gagagagaaa teggagagaa eeggagagaga ceeggagaaee ceeggeeaaa eeggagagaga eegaaceaee ceeggeeeaaee ceeggeeeaaee ceeggagagaa eegaacaaaee cacaaaaee aaaaaataage 58600 gagaacaaaee gagaacaaaa tegaagagaga gagaacaaaa aaaaaataa aaaaaataa aaaaaataa aaaaataa aaaaaa	ggtctactgc	cgtcttccgt	gggctgtttc	agttcctatt	acccagagte	aaactcccag	57840
cactitigeat tictigicos atalicitagi coatigagat acatiticatig tettigagic 57960 teacacaaga tiggigea teacacaagagi tettifiaa attitititit tigaaga geacacaca caccacacacacacacacacacacacacaca	cattccctac	ctgattccag	acttggagtc	cagagettta	acctcttcag	gccaactccc	57900
tacttacaaa gtaaattttg ctgttttta atttttttt tagaatgaga tcttgccctg 58020 tcaccaaggc tgtggtgcaa tgacgcaac tcggccact ggaaccaccc gcacctccgg 58080 tcaagcgata tttttttat cagccccaa gtagctgtga ttacagacag gcaccaccac 58140 gcccagctaa tttttttat cttttagtaa agacagggtt tcaccatgtt ggccaggctg 58200 gcgtgagcc actgtgccca cttttagtaa agacagggtt tcaccatgtt ggccaaggctg 58200 gcggtgagac actgtgccca actgtgccca actgtgccca actgtgccca actgtgccca actgtgccca actgtgccca actgtgccac actgtgccca actgtgccca actgtgccca actgtgccac acattttg ccaaatatta gaccaaaaaa atctcaaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa acaacttaaa caaaattaaag 58300 tactcaagtc aaccccagaa cttttgggagg ccaaggggg cagaaccacttaa caaaattaaa ggccaagaac tccaaggccgg 58500 ccaaggaccag cctgggcaat tgggggagg ccacaattata gaccaaaaaa atctcaaaaa 58300 tactcaagtc aaccccagaa cttttgggagg ccaaatataa caaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac ggcggtggc 58500 ccaaggaacac ccggccaaa tggcaagaag ggcaaaaaatac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac aaaaattaac ggaagaacacac tccaacacac ccaagaccac ccaagaccac ccaagaccacac ccaagaccacacaca	cactttgcat	ttctgtccct	atatcttagt	ccatggagat	acatttcatg	tctttqaqtc	57960
tcaaccaggc tgtggtgaa tgacgccate tgggtcact gcaacctccg cctcctggt 58080 tcaagcgatt catctgcct agcetcccaa gtagctgtga ttacagacag gcaccaccac 58140 gcccagctaa tttttttat cttttagtag agacagggtt tcaccatgtt ggccaggctg 58200 gtcttgaatt cctgacctcg tgatetgcc atctcggct cccaaagtgc tgaggattaca 58380 atatagaaag aaaatccaaa aaatccaaa aaatccaaa aaccaaaaa atctcaaaa atctcaaaa aatccaaa aaatccaaa aaatccaaa aacccccaatta taccactg acccccaatta aaataaaatt agaaattaac cacaacttaa accacagcc 58440 atactcactg acccccaata aaataaaatt agaaattaac cacaacttaa agagaggaa gaggaaataa acatcaaaat tacaaaagtc aacacccag cttgggcaat atggtgaaac cccgtttca ctataggtga gggtgggc gagagaatca ccgggaaccag cctggcaat atggtgaaac ccacttggga ggccaacacac ggtcaggaga ttgcaggaga gagaccaaca ccacatttaa acaaaatta agaaattaac cacaacttaa agaaattaac cacaacttaa agaaattaac cacaacttaa agagagaaa gaggaaataa acatcaaaat tacaaaagtc gggtggggc caggggggg ccaaagggggg cagaaccaac gggtcgagaac gggtgggaga gggggggggg	tacttacaaa	gtaaattttg	ctgtttttta	atttttttt	tgagatggag	tettacceta	58020
ccaagcataa ttttttat ctttagtag agacagggtt tcaccatgtt ggccaggctgggtggtcttgaatt cctgacctcg tgatctgcc actctggcct cccaaagtgc tgagctggggcc actgtggcca gccaattttg cttttattat atttcattgc tgagctggggcc actgtggcca dattgcata ccaaaattata gaccaaaaaa attccaaaa aaataccaaa aagcaacaaat ctccaaaaa aaataccaaa aagcacaaata cccaagaccgctgagacca ccccaataa aaataccaaa aagcacaaaaa atcccaaaa aaataccaaa aagcaacaaat cccaaaaattaccaaca aagcagggggg caaaataaagcagggggggggg	tcacccaggc	tgtggtgcaa	tgacgccatc	tcggctcact	gcaacctccg	cctcctgggt	58080
gcctagata ttttttat ctttagag agacagggt tcacatgtt ggcaggctg 58200 ggctgagcc actggccca tgatctgcc atcteggcct cccaaagtgc tgagattaca 58260 ggcgtgagcc actggccca gccaaatttt gccaaagtgc tgagattaca 58260 gaggataagt ttacagtgct atatgcatt cccaaaaataa atctcaaaa aaatccaaa aaatccaaa aaatccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa aaataccaaa caaaataaag 58500 tcacgcctgt aatcccagca ctttgggag ccaaggcggg caaggcgggg caaggcataagt ccgggcaata tacaaaatt agaaattaac caaaataaag 58500 tcacgcctgt aatcccagca ctttgggag ccaaggcggg caaggcgggg caaggcatagt gatgtgtgc tgtaatccag cccgtttcca ccaaaataa aaaataaaa 58600 tcacggaaccag gatgtggaaa cccgtttcca ccaaaaataa aaaaataagc gagcaaaaca tttggagagaa gagaacaaa cccgtttcca caaaaataa aaaaataagc gagcaacaa tggtgagaac cccgtttcca ccaaaaataa ggcaggaga 58600 tcagaccaagg gagacaaaa atggtgagaac cccgtttcca ccaaaaataa ggcaggaga 58600 tcaagcaaagc gagaccaaag gagacaaaaa aaaaaaaa	tcaagcgatt	catctgcctc	agcctcccaa	gtagctgtga	ttacagacag	gcaccaccac	58140
gtcttgaatt cctgacctcg tgatctgccc atcteggct cccaaagtgc tgagattaca 58260 ggcgtgaggc actgtgcca gccaattttg cttttttat atttttatat gaccaaaaa atctccaaaa 58380 aattagaaag aaaatccaaa aaatccaaa aaataccaaa aaataccaaa aagcaacaat ctcaaaaag 58500 aattagaaccagc cttttgggagg ccaaggcgggg cagatcaaa ggcgggtggc 58560 cacggctgt aatcccagac cttttgggagg ccaaggcggg cagatcacaa ggcgggtggc 58560 cacggcatag gatgacaaa cctttaggaaggaaaaa cctttaggaaggaaaaa cctttaggaaggaaaaa cctttaggaaggaaaaa cctttaggaaggaaaaa cctttaggaaggaaaaaaaaaa	gcccagctaa	tttttttat	cttttagtag	agacagggtt	tcaccatqtt	ggccaggctg	58200
ggggtaggcc actgtgcca gccaattttg cttttttat atttcattgc tatatgtta 58320 gaggataagt ttacagtgct atatgcatc ccaaatatta gaccaaaaa atccaaaa caataccaaa aaatccaaaa aaatccaaaa aaatccaaaa aaatccaaaa aaatccaaaa caacacattaa ccaaacttaa ccaaaatacagcc 58440 acaccagcctg aatcccaaga gaggaaaataa acatcaaaaat tacaaagtct agggggggg 58500 tcacggcctgt aatcccagca ctttgggagg ccaagagggggg ccaagagggggg caagactacaa ggtcaggaat 58620 tcagggaccag cctggccaat atggtgaaac cccgtttcca ctaaaaaatac aaaaattagc 58680 caggcatagt gatgtgtgc tgtaatccag ccacttggga ggctgaggca ggagaatcac 58740 tgaaaccaagg gagaccaaag ttgcagtggg ccaaaaatcgt ggccactgac ttcggcctgg 58800 ttggtgtttt tttttgagac ggagtctcac tcaaaaaaa aaaaaattac aaactcttta gatagaaatt 58860 ttggtgtttt tttttgagac ggagtctcac tctgtcgcag aggctggagt gcagtggggc 58920 caagataggtg tttcaacat gggccaacaa taggctgagt gaggtggggc 58920 acatgcact aggattacag gggccaacaa taggccaacaa taccaagag ccaagtgccc ccaagtagcc ccaactagag ccaagacagac ccaagtagcc ccaagtagcc ccaactagag ccaagacagacc ttttggcagg tggctcaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaa	gtcttgaatt	cctgacctcg	tgatctgccc	atctcggcct	cccaaagtgc	tgagattaca	58260
gaggataagt ttacagtget atatgcatte ccaaatatta gaccaaaaaa atetecaaaa 58380 aaattegaaa aaaatecaaa aaataccaaa aaataccaaa aagcaacaat ctacacagace 58440 atateteagte acceccaata aaataaaat aaaatacaaaa cacaacttaa caaaataaaag 58500 tacteaget agagaggaaa gaggaaataa ccacaaattaa caaaataaaag 58500 tacteaget aateecagea ctttgggagg ccaaaggeggg cagatcacaa ggtcaggagt 58600 teaggacaag catggegaaa ateggtgaaac ccegttteca ctaaaaataa aaaaattage 58680 caaggaataag gatggtgee tgaateeag ggetgaggaa ggagaacaac 58740 tgaaccaagg gagaccaaga ttgcagtgag ccaaaaataa aaaaattaac aaaaattage 58680 gtgacaaaage gagactecat tttttgagac ccaacttggga gecactgcac tteggeetgg 58800 gtgacaaaage gagactecat ctcaaaaaaa aaaaaattac aaactettta gatagaaatt 58860 ttggtgttt tttttgagac gaggteteac tetgteegag aggetggagt gcagtgggac 58920 tatgteaget tttcacaget cacctecet ggattecaag ctagtttta tatttttagt 59040 agagatggtg tttcaccaag ttggecaage tggeteaaa ctcctgacet caagtggtee ccaacaaggaggagaatecac tggeteaaaa taggeggggg ctgagggagaa aaagaaaaaa aaaaaaaaa aaaaaaaa	ggcgtgagcc	actgtgccca	gccaattttg	ctttttttat	atttcattgc	tatatqttta	58320
atattagaaag aaaatccaaa aaattaccaaa aaattaccaaa aagcaacaat ctcacagacc sactcacactcactcactcactcactcactcactcactca	gaggataagt	ttacagtgct	atatgcattc	ccaaatatta	gaccaaaaaa	atctccaaaa	58380
tactcaagtc agagggaaa gaggaaataa acatcaaaat tacaaagtct agagggggggggg	aattagaaag	aaaatccaaa	aaatctcaaa	aaataccaaa	aagcaacaat	ctcacagacc	58440
tactcaagtc agagaggaa gaggaaataa acatcaaaat tacaaagtct aggcggtggc 58560 tcacgcctgt aatccagca ctttgggagg ccaaggcggg cagatcacaa ggtcaggaat 58620 ccggataagt gatgtgtgc tgtaatccag ccacttggga ggctgaggca ggagaatcac ggagaatcac ggagaaccagggggggggg	atactcactg	acccccaata	aaataaaatt	agaaattaac	cacaacttaa	caaaataaaq	
tcacgcctgt tcgagaccag tcgagaccag tcgagaccag tcgagaccag tcgagaccag tcgagaccag tcgagaccag tcgagaccag tagtgtgcc tgtaatccag tagtgtgcc tgtaatccag tagtgtgaaac tcacattgga gagactcaat ttgcagtgaa ttgcagtgag tcaaaaattac tagtgaaacac tggctaagag gagactcact tgtaatccag gagactcact ttgagag gagactcact tttttgagac tcaaaaaattac tactgtgag gagactcact ttttttgagac gagactcact tcttttagacc tcaactgtgag gagactcact tttttttagacc tcaactgtagag gagactcacc tcaacaccag gagactcacc tcaacaccag gagactcacc tcaacaccag gagactcacc tcaacaccag gagactcacc tcaacaccag gagactcacc tccaacaccag gagactcacc tccaacaccag gagaccacca gagaccaccag gagactcacc tccaacaccag gagaccaccag tcaacttcag gagaccaccag accaccag accaccag accaccag gagaccaccag tcaacaccag accaccag accaccag gagaccaccag tcaacaccag accaccag accacaccag accacaccag accacaccag accacaccac tcaacaccaccaccaccaccaccaccaccaccaccaccac	tactcaagtc	agagaggaaa	gaggaaataa	acatcaaaat	tacaaaqtct	aggcagtagc	58560
ccagaccag cctggcaat atggtgaaac cccgtttcca ctaaaaatac aaaaattagc 58680 caggcatagt gatgtgtgcc tgtaatccag ccacttggga ggctgaggca ggagaatcac 58740 ggcacaaagc gagacgaaga ttgcagtgag ccaaaatcgt gccactgcac ttccggcctgg 58800 ctcaaaaaac aaaaattac aaactcttta gatagaaatt 58860 gtgacaaagc caccgcaacc tccaaaaaaa aaaaaattac aaactcttta gatagaaatt 58860 ggagtctagct caccgcaacc tccaacaca aggctggagt gcagtgggac 58920 tatgtcagt caccgcaacc tccaactcct ggagtccaa aggctggagt gcagtgggac 58980 ccaagaatggtg tttcaccatg ttggccagg tggtccaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaa gtgctcagat tacaggcgtg agccaccgca cccacctag 59100 acctgcttca gcctccaaa tggctcagat tacaggcgtg agccaccgca cccacctag 59100 agagctgagg cgtgggaga tcacttgggc ccaggagttc aggaccagca tgggtgacag 59280 agaacagacc tgtctctatt tatttgaaaa aaaaaaaaaa	tcacgcctgt	aatcccagca	ctttgggagg	ccaaqqcqqq	cagatcacaa	ggtcaggaat	
caggcatagt gatgtgtcc tgtaatccag ccacttgga ggctgaggca ggagaatcac 58740 tgaacccagg gagacgaaga ttgcagtgag ccaaaaatcgt gccactgcac ttcggcctgg 58800 gtgacaaagc gagactccat ctcaaaaaaa aaaaaattac aaactcttta gatagaaatt 58860 ttggtgttt tttttgagac ggagtctcac tctgtcgcag aggctggagt gcagtgggac 58920 ccaagtagct caccgcaacc tccatctct ggattcaagc aattctcctg tctcagcctc 58980 gcgccaacca aggattacag attctcctg tctcagcctc 58980 gcgccaacca aggattacag ctagtttta tatttttagt 59040 aggagttggt tttcaccatg ttggccaggc tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaaa tggctcacga tggctcacga cccaccaa ggggcacaa tggctcacgc ctgtaatctc agcacttcag 59160 acctgcttca caacatgagg ccagggcacaa tggctcacgc ctgtaatctc agcacttcag 59220 ccaggagtcacagaccacaga caacagaccag tgtctctatt tatttgaaaa aaaaaaaaaa	tcgagaccag	cctggccaat	atggtgaaac	cccqtttcca	ctaaaaatac	aaaaattagc	
gaacccagg gagactcat ttgcagtgag ccaaaatcgt gccactgcac ttcggcctgg 58800 gtgacaaagc gagactcat ctcaaaaaaa aaaaaattac aaactcttta gatagaaatt 58860 ttggtgttt tttttgagac ggagtctcac tctgtcgcag aggctggagt gcagtgggac 58920 tatgtcagct caccgcaacc tccatctct ggattcaagc aattctcctg tctcagcctc 58980 aggagtgggt gtggactgggac ggagtctcac tccatcagccag ctagtttta tatttttagt 59040 aggagttggt tttcaccatg ttggccaggc tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaaa tggctcacgc tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaaa tggctcacgc ctgtaatctc agcacttcag 5920 tatagaaattt gaggctgagg tcaccaggc tggccaccac cccacccag 59160 acctgctcagg cggggagga tcacttgggc ccaggagttc aggaccagca tgggtgacag 59280 aggacagaccc tgtctctat tattgaaaa aaaaaaaaaa	caggcatagt	gatgtgtgcc	tgtaatccag	ccacttggga	ggctgaggca	ggagaatcac	
ggactaaagg gagactccat ctcaaaaaa aaaaaatac aaactctta gatagaaat 58860 ttggtgtttt tttttgagac ggggtcacac tccatctcct ggattcaagc aattctcctg tctcagcct 58980 ccaagtagct aggattacag gcgccacca cccaagacccag ctagtttta tatttttagt 59040 acctgcttca gcctccaaa ttggccagc tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaa ttggccagc tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaa tggctcagat tacaggggtg agccaccgca cccaacctag 59160 acactgcttca gcctccaaa tggctcacgc ctgtaatctc agcacttcag 59220 cgtgggagga tcacttgggc ccaggagttc aggaccacgca tgggtgacag 59280 agacagaccc tgtctctat tattgaaaa aaaaaaaaaa	tgaacccagg	gagacgaaga	ttqcaqtqaq	ccaaaatcgt	gccactgcac	ttcaacctaa	
ttggtgtttt ttttgagac ggagtctcac tctgtcgcag aggctggagt gcagtgggac 58920 tatgtcagct caccgcaace tccatctct ggattcaage aattctcctg tctcagctc 58980 ccaagtagct aggattacag gcgccacca ccagacccag ctagtttta tattttagt 59040 acctgcttca gcctccaaa ttggccaggt tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaa ttggccaggt agccaccgca cccaactag 59160 acaggagagagagagagagagagagagagagagagagag	gtgacaaagc	gagactccat	ctcaaaaaaa	aaaaaattac	aaactcttta	gatagaaatt	
caagtaget cacegeaace tecateteet ggatteaage aatteteetg teteageete 58980 ceaagtaget aggattacag gegeeaace ceagaceag etagttitta tattittagt 59040 aggattggtg titeaceage tiggeeage tiggeeage tiggeeage eeceaceage 59160 acetgettea geeteeaaa teggeeage tiggeeage tiggeeage eeceaceage 59220 gaggetgagg eggggagga teaetiggg eeaaaaaaaaa aaagaagaga agaaagaaaa 59340 titeaacatga aaagtatee teaaaceet egagatgii geaaaaageg agaaagaaaa 59340 titeaacatga aaagtatee teaaaceet egagatgii geaaaaageg agaaagaaa 59400 aaaceetgiaa teeceaacat etggegagee gaataagaaa gaggeegggi giggii gaggatea agageeggii giggii gaggagaa 59280 agaceatee tiggii gaaaacee tiggii geaaaaagaa agageegggii giggii gagaaaacee tiggii gagaaaacee tiggii gaaaacee tiggii gaaacee tiggii gagaaacee tiggii gagaaacee tiggii gaaacee tiggii gagaaacee tiggii	ttggtgtttt	tttttgagac	ggagtctcac	tctqtcqcaq	aggetggagt	gcagtgggac	
ccaagtagct aggattacag gcgccacca ccagacccag ctagtttta tatttttagt 59040 agagatggtg tttcaccatg ttggccaggc tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaa gtgctcagat tacaggcgtg agccaccgca cccacctag 59160 atagaaattt caacatgagg ccgggcacaa tggctcacgc ctgtaatctc agcacttcag 59220 agagctgagg cgtgggagga tcacttgggc ccaggagttc aggaccagca tgggtgacag 59280 agacaagacc tgtctctatt tatttgaaaa aaaaaaaaa aaagagagag agaaagaa	tatgtcagct	caccgcaacc	tccatctcct	ggattcaagc	aattctccto	totcagooto	
agagatggtg titcaccatg ttggccaggc tggtctcaaa ctcctgacct caagtgatcc 59100 acctgcttca gcctccaaa gtgctcagat tacaggggtg agccaccgca cccaacctag 59160 atagaaattt caacatgagg ccaggacaaa tggctcacgc ctgtaatctc agcacttcag 59220 agaccagacc tgtctctatt tatttgaaaa aaaaaaaaa aaagagagag agaaagaa	ccaagtaget	aggattacag	gcgcccacca	ccagacccag	ctagttttta	tatttttagt	
acctgettea geetecaaa gtgeteagat tacaggegtg agceacegea ecceacetag 59160 atagaaattt caacatgagg eegggeacaa tggeteaceg etgtaatete agcaetteag 59220 gaggetgagg etgeggagga teaettggge ecaggagtte aggaecagea tgggtgacag 59280 agacagacae tgtetetat tatttgaaaa aaaaaaaaaa aaagagagag agaaagaa	agagatggtg	tttcaccato	ttggccaggc	tootctcaaa	ctcctgacct	caagtgatcc	
atagaaattt caacatgagg cegggcacaa tggctcacgc ctgtaatctc agcacttcag 59220 gaggctgagg tcgtgggagga tcacttgggc ccaggagttc aggaccagca tgggtgacag 59280 agacagaccc tgtctctatt tatttgaaaa aaaaaaaaaa	acctacttca	gcctcccaaa	gtgctcagat	tacaggcgtg	adccaccaca	ccccacctag	
gaggctgagg cgtgggagga tcacttggg ccaggagtt aggaccagca tgggtgacag 59280 agacagacce tgtctctatt tatttgaaaa aaaaaaaaa aaaagaggag agaaagaa	atagaaattt	caacatgagg	ccgggcacaa	taactcacac	ctgtaatctc	agracttrag	
agacagaccc tgtctctatt tatttgaaaa aaaaaaaaaa	gaggetgagg	cataggagga	tcacttoggc	ccaggagttc	addaccadca	tacatacca	
ttcaacatga aaagtatctc tcaaaccctt cgagatgttg gcaaaaaagcg actcaaagga 59400 aaatgtatta ctgtgtgtga atttgcttga aaataagaaa gaggccgggt gtggtggcta 59460 acacctgtaa tcccaacact ctgggagtcc gaatcaagtg gatcatgagg tcaggagatc 59520 gagaccatcc tggctaacat ggtgaaaccc tgtctctact aaaaatacaa aaaattagct 59580 aggcgcggtg gctcatgcct gtaatcccag cactttggga ggctgaggca ggtggatcac 59640 ctgaggtcag gggtttgaga ccagcctggc ctacatggtg aaacctcgtc tcttcacaa 59700 atacaaaaat tagctgggcg tggtggtgg tgcctgtaat cccagctact cagaggctga 59760 ggcaggagaa tcgcttgaac ccgggaqqcq gaggttgggg tgagccgaga tcgccact 59820	agacagaccc	totctctatt	tatttgaaaa	22222222	aggaccagca	agaagaaat	
adatgtatta ctgtgtgtga atttgcttga aaataagaaa gaggccgggt gtggtggcta 59460 acacctgtaa tcccaacact ctgggagtcc gaatcaagtg gatcatgagg tcaggagatc 59520 gagaccatc tggctaacat ggtgaaaccc tgtctctact aaaaatacaa aaaattagct 59580 aggcggggtg gctcatgcct gtaatcccag cactttggga ggctgaggca ggtggatcac 59640 ctgaggtcag gggtttgaga ccagcctggc ctacatggtg aaacctcgtc tcttcacaa 59700 atacaaaaaat tagctgggcg tggtggtggg tgcctgtaat cccagctact cagaggctga 59760 ggcaggagaa tcgcttgaac ccgggaqqcq gaggttgcgg tgagccgaga tcgcaccact 59820	ttcaacatga	agagtatete	tcaaaccctt	casastatta	aaagagagag	agaaagaaac	
acacetglaa teccaacact etgggagtee gaatcaagtg gatcatgagg teaggagate 59520 gagaceatee tggetaacat ggtgaaacee tgtetetaet aaaaatacaa aaaattaget 59580 aggegeggt geteatgeet gtaateecag caetttggga ggetgaggea ggtggateae 59640 etgaggteag gggtttgaga ceagcetgge etacatggtg aaacetegte tettetacaa 59700 atacaaaaaat tagetgggeg tggtggtggg tgeetgtaat eecagetaet cagaggetga 59760 ggeaggagaa tegetgaac eegggaqqeq gaggttgagg tgageegaga tegeaceact 59820	aaatgtatta	ctatatataa	atttactta	asatasassa	gcaaaaagcg	acccaaagga	
gagaccatec tggctaacat ggtgaaaccc tgtctctact aaaaatacaa aaaattagct 59580 aggcgcggtg gctcatgcct gtaatcccag cactttggga ggctgaggca ggtggatcac 59640 ctgaggtcag gggtttgaga ccagcctggc ctacatggtg aaacctcgtc tcttctacaa 59700 atacaaaaaat tagctgggcg tggtggtggg tgcctgtaat cccagctact cagaggctga 59760 ggcaggagaa tcgcttgaac ccgggaqqcq gaggttgggg tgagccgaga tcgcaccact 59820	acacctotaa	toccaacact	ctagaactaa	gaatgaagaaa	gayyccyygt	tasaasaasa	
aggogoggtg geteatgeet gtaateceag eactitiggga ggetgaggea ggtggateae 59640 etgaggteag gggtttgaga eeageetgge etacatiggtg aaacetegte tettetaeaa 59700 atacaaaaat tagetgggeg tggtggtggg tgeetgtaat eeeagetaet eagaggetga 59760 ggeaggagaa tegettgaae eegggaqqeq gaggttgegg taageegaga tegeaceaet 59820	gagaccatco	tooctaacact	catassace	tatatatat	garcargagg	ccaggagate	
ctgaggtcag gggtttgaga ccagcctggc ctacatggtg aaacctcgtc tettetacaa 59700 atacaaaaat tagetgggcg tggtggtggg tgeetgtaat cecagetaet cagaggetga 59760 ggcaggagaa tegettgaac cegggaggeg gaggttgegg tgageegaga tegeaceaet 59820	aggegeate	actcatacat	grantage	cyclictact	addadaCaCaa	adaattagct	
atacaaaaat tagctgggcg tggtggtggg tgcctgtaat cccagctact cagaggctga 59760 ggcaggagaa tcgcttgaac ccgggaqqcq qaqqttqcqq tgaqccqaqa tcgcaccact 59820	ctdacctcac	gaatttaaa	graarcccag	cactttggga	ggctgaggca	ggtggatcac	
ggcaggagaa tegettgaac eegggaggeg qaqqttqeqq tqaqeeqaqa tegeaccaet 59820	atacasasat	tacetecese	tootootoo	ccacatggtg	aaacctcgtc	tettetacaa	
ggcaggaga tegettgaat tegggaggeg gaggttgegg tgageegaga tegeaceaet 59820 acaetecage etgggeaaca geetgggtga cacagtgaga etceatetea aaaaatacaa 59880	acacadadat	teattance	caggaggg	rgcctgtaat	cccagctact	cagaggctga	
acacccage cogggeadea georgggiga cacagigaga ciccatetea aaaaatacaa 59880	ageayyayaa	ctccccaac	ccssgaggeg	gaggttgcgg	tgagccgaga	tcgcaccact	
	acacccage	CCGGGCaaca	gcccgggcga	cacagugaga	CECCAECECA	aaaaatacaa	22880

33/122

59940 aaaattaget gggtgtggtg geetgegeet gtagteeeag etaeceggga ggetgaggea ggagaatgga gtgaacctgg gaggaggagc ttgcagtgag ccgagatccc accactgcac tccagcctgg gcgacagagc aagactcttg tctcaaaaaa aagaaaaaaa aaggaaaaaa 60000 60060 gaaccetgat aataaagaaa ccaaatgtte aacteteaaa geteggacae tttaaagaaa 60120 taattaataa aggcagaagt taaagggagg atgataaagc aatttttttt gttggttttt 60180 ttgagatgga gtcttgctct gtcacccagg ctggagtgca gtgatgcgat cttggctcac tgcaacctct gcctccggg ttcaagcaat tctcctgcct cagcctcctg agtagctggt actacaggtg cgcgccacct ggcccagcta atttttgtat ttttattaga gacggggttt 60240 60300 60360 60420 caccatattt gttaggctgg tctcaaactc ctgatctcag gtaatctgcc cacctcggcc teteaaagtg etgggattae aggeaggege eacegegeet ggeetaaage aaaatattgg 60480 ttetgtgeaa aaggteaata aaaagageaa aegtttacaa aetggageea geaceeatte 60540 ageteagtgt gtetggagaa aaaacaatet egetteagaa tteatgatta egeageeett 60600 titgetieet aaaaateeta etatgttget gitgaceatt etetetett etetetetet 60660 tgctttctct ccagaaaagc tattcagaca ttctcctctt tcctcaaacc tccaacactt 60720 cctcctccat ccttagcctc agctgctgac ctcacttcta atcattgaga aaccaggaga agcatttaag agtgaacctc cgcctccccg cacgggcaaa accacccacc cacagaattg 60780 60840 tgececaatt etgegteete teeteteace atggatggae ggteeagget eegageeaaa 60900 gecaggeete ceetggaget etggatecae cacetgeage tteteaggea gggeeceage 60960 ageteceetg etecettgta ceateaatee eteceeteae tgggteaete eeaacaatat 61020 atatatttag tgatgtttct cccatgtggt aaaatcactt agcctctctc ctcccccagc 61080 tactatecta titgittett tecatietet geaaaaette teaaageatt gigtetaigt 61140 getgaeteca tttatettet ecegttetet getgagteet teccaeagae teteaececa 61200 gttactccat gaaatgacct ctgcactgcc acatecaatg gtgaatgite agttettaat tttattcagt ettteageag eatttgaeet ggeegateae teeetettet taaaaataet 61260 61320 tttctcagcc aggcgtgatg gctcacacct gtaatcccaa cactttggga ggccaaggcg 61380 61440 ggaggatcat gagagcccag gagttcaaga tcagcctggg caacatggca agaccctatc tctacaaaaa ctaaaaagta gccagtgtga tggcatgcac ctgtagtccc atctacttag 61500 gaggetgagg cagtaggatg acttgageet gggaaalcaa ggetgeagtg agecatgati 61560 gcaccactgc actocagoot gagtgacago gagaccotgt otcaaaaaga caaaatagga 61620 aacttttete agcatattee tetgattete etgetgette tgtetgeaca gatteagtet eetttgeegg ttetteetea teeteetgat etettgacet tgaagtgeee eagagtacag 61680 61740 tettttttt ttttttgag aegeagtete gtetgteace caagetggag tgeaatggeg 61800 aggteteage teatgeaace tetgeeteet gggtteaage gatteteetg ceteageete 61860 ccaagtagcc aggactacag gcacatgcca ccatgcccag caaattgttg tatttttagt 61920 agagacaggg ttttactata ttggccaegc tggtetcaaa eteetgaact egtgaaccae 61980 62040 eegeetegge eteccaaagt getgagatta caggeatgag ecaccacace eggeecagag tacagtetit agaeggeete tetacetata ettgeteece teataaacte etectgeete 62100 atggetttaa ataceategg tagaetgatg acteecatat ttetetttt tttttggaga 62160 eggagteteg etcagteeec caggetggag tgeagtggeg egatetegge teaetgeaag 62220 ctccacctgc caagttcaca ccattctcct acctcagcct ctccagtage tgggactaca 62280 ggcacccgcc accacgcctg gctaattttt ttgtattttt agtagagatg gggtttcacc 62340 atgttageca ggatggtete gatetéetga cetegtgate egeceatete ggeeteceaa 62400 agigctggga tiataggtgt gagccaccgt gcccagccga tgactcccat atttctatct cttgctgtgt gggagttctc ctcagaactc catactcata aatccaactc tcataaatag 62460 62520 tateteaaat gggeaatatg eteaaaagte aatteetaet titeteeeta aaettgetit eetgeagtet eeaceatett aatgteeaat etaacattag gaggeaaaaa etitgaagte attettgaet etteteatt acacaceeta teeaatetti etgeagatee agtegaeeee 62580 62640 62700 caaatccagt tageteteat cateteeeet gttaceeet ggteeaggee atetteetet 62760 cteacetgaa teactgeage attetectea etggtetett tggttetgtt tteacteeae 62820 cttagcatag tetecacaga geagteagag ggateetttt aaagtgtaat teecateetg teectgetet geteaaaace etgtegtgat teeegtttta atetgteaga ttaaaageea 62880 62940 gagtetttee agtgacetae atgatetgee tattateace teccaettet tteccettge 63000 teactecact ceagetetge agetgteett tetgttteet gaacageeca gattttgett etttagaace tttgtatttg etgteceete tgtetggaat gttttteeag gaagteacet 63060 63120 63180 ggctctctcc tgcacttcct tcctgaccac catgtttaaa aatcactcaa acacacttca ggccggacat ggtggctcac gcctgtaatc ccagcacttt gggaggccaa ggtgggtgga 63240 teacetgagg teaggagtte gagaceagee tggccaacat ggtgaaactt egtetetact 63300 acaaatacaa atagtagcca ggtgtagtgg cacacacctg taatctcagc tactcaggag gctgaggcag gagaatcgct tgaacccaga aggcagagga ggtgcagtga gccaagatca 63360 63420 cgccacaaca ccccagcctg ggtgacagag caagacccca tctcaaaaaa aaaaaaagaa 63480

aaaaaaatca	cacaaacaca	cttctcttca	tattcctttt	ccaagtttta	tttttctcca	63540
gaatacttta	cattgtttta	atggaagttc	tccgtttccc	cccaactaga	atggatactt	63600
			caagtactaa			63660
			cagggtggta			63720
			cttttgttcc			63780
ctacttotac	ataataagca	ttcaataaat	atttgttgaa	tgaatgactt	gttgaatgaa	63840
ttaatctcag	aaatgcagga	ctggttctac	attagaaaat	ttttcaaggt	cattetetet	63900
tgtcgtaaca	cattaagaga	ggaaaatttt	gtactctaaa	tcatttgata	aaatacatac	63960
tgatttctgt	tttcaaaaac	tcttagtggc	tgggcgaggt	goctcacate	tataatccca	64020
gcattttggg	aggacgaggt	gggcggatca	cttgaggtca	ggagtttgag	accagected	64080
ccatcatggt	gaaaccctat	ctctactgaa	aatagaaaaa	ttagccgggt	ataataacac	64140
			gaggcaggag			64200
			ttgcactcca			64260
			ttaggaatcc			64320
			taaaccttaa			64380
						64440
			atgcctgcaa			64500
			aaaaaataaa			64560
cagaggergg	cagcaaggac	tosastass	tacagtactt	gcccggagca	ggcggatggc	64620
tacacccccg	ttaatataa	ccagetgget	gggggacgct	ccagugugug	agragga	64680
			actggggaga			
			ttctgtcaaa			64740
			aggtgagaga			64800
caccacaget	geggaggegg	acagegacag	tggtgggccc	tggccagaet	tttcatgete	64860
aaaggtggtg	grigitette	ctacttettg	tccctccagg	gerreerreg	cccgcgcgcc	64920
			ttttaaattt			64980
			gaaccctgct			65040
			aaagtggcaa			65100
cagtgactgg	atcatgagag	ctctgacctc	atgagtggat	taatggatta	atgagttgtc	65160
argggagrgg	catcagtggc	tttataagag	gaagaattaa	gacctgagct	agcatggtcg	65220
cccctcacc	atttgatate	ttacactgcc	taggggctct	gcagagagtc	cccaccaaca	65280
			ccttgtactt			65340
gaaataaatg	CCTTTTCTT	atgaattacc	cagtttcaga	tattctgtta	taaacaatag	65400
aaaacgaact	aaggcaaact	ctcatgattc	tactgccatg	ccattccaat	aaactccctt	65460
tatgettaag	agagccagag	ttggccaggc	gtggtgactc	acgcctgtaa	ttccagcact	65520
ttgggaggee	gaggcaggtg	gatcacaagg	tcaggagatc	gagaccatcc	tggctaacac	65580
			aaaaattagc			65640
tgtagtccca	gctactcggg	aggctgaagc	aggaggagaa	tggcgtggac	ccaggaggcg	65700
			gcactccagc			65760
			atttctgttg			65820
			caatttaaag			65880
			acattcattc			65940
			gggccgggtg			66000
cctagcactt	rgggaggccg	aggcaggtag	atcacctgag	gccaggagtt	cgagaccaac	66060
			taaaaataca			66120
			agactgaggc			66180
ggaggcagag	gttgcagtga	gccgagatcg	caccactgca	ctccagcctg	ggcaacaaga	66240
gcgaaactcc	acctcgaaaa	aaaaaaaaa	aaaaaaagag	ggccggggct	gggcgcagtg	66300
geteaegeet	gtaatcccag	cactetggga	ggccaaggca	ggagaattac	gaggtcagca	66360
gatcgagacc	agcctgacca	acatggtgaa	accccatctc	tactaaaaat	acaaaaatta	66420
rccgggcgtg	gtggcgcaca	cctctagtcc	cagctacttg	ggaggctgag	gcaggagaat	66480
cgcttgaacc	cgggaggcag	aggttgcagt	gagccgaaat	catgccactg	cactccagcc	66540
tgggtgacag	agtgagactc	cgtctcaaaa	aaaaaataaa	aaaaaaaaa	gaattcaaaa	66600
attgtagagt	tatagtgtgc	ttctagttta	gttgagagga	catctgtcct	tcaaggaagg	66660
			atcaatccag			66720
cgatcacage	agtaagatag	gaagagcacc	tttgtacatt	tagctcatgt	tgagataagc	66780
cactgacaga	gctgaaggaa	gctcacagtt	ctgggttcca	tcctttggca	tttaaaaaga	66840
aaagtgctaa	gaaaattcgg	ttggtcacgg	tggctcacgc	ctgtaatccc	aacactttga	66900
gaggccaagg	caggcagatc	acgaggtcag	gagttcgaaa	ccaqcctqqc	caacatqqtq	66960
aaaccccgtc	tctactaaaa	acagaaaaat	tagccgggca	tggtggcgca	tgcctataat	67020
cccagctact	caggaggctg	aggcaggaga	attgcttgaa	cccgggaggg	ggaggttgca	67080

gcgagtgaga	gcaggccact	gcactccagc	ctgggagaca	gagcaagact	ctgtctcaaa	67140
aaaaaaaaag	aaaaaaagaa	agaaaggaaa	aaaagaaaga	aaaaaaaaga	aaaaagaaaa	67200
ttcaggccag	gccaggcctg	gtggctcaca	cctgtaatcc	caacactttg	ggaggctgaa	67260
				agcaacatag		67320
				ctcacgcctg		67380
				atggagacca		67440
				aaccgggcgt		67500
				tggcgtgaac		67560
				ctgggagaga		67620
				gtgtgctggc		67680
				gagcccagga		67740
				gaccctctct		67800
				aaggaaggaa		67860 67920
				aagaaaagag ttatttattt		67980
				tgcagtgatg		68040
				ctcagcctcc		68100
				attctttgta		68160
				aagccatcca		68220
				ctaatcgggt		68280
				agtggctgtt		68340
				tccagagtag		68400
				ttatattgag		68460
ccacctagag	gatatatgta	gtaccgaact	gtgtgcgcag	ggaggctgag	gttgcagtga	68520
				caagatttca		68580
aaaaaaaaa	aaaaaaaaa	aagaattgaa	agtaaggtct	tgaagagata	tttgtgcctg	68640
tatggtcata	gcagtattaa	ctttgaccca	ctagctaaaa	cacaaaagca	acatgtgtct	68700
				acaattgaat		68760
ctttaaaaag	gaataaaagg	ctggatgcgg	gggctcacgc	ctgtaatcct	aacactttgg	68820
gagactgagg	tgggtggatc	acccgaggtt	aggagtttga	gaacagcctg	gccaacatgg	68880
tgaaacttca	tctctactaa	aaatactaaa	attagccggg	catggtggca	cttgtctgta	68940
				aactcaggag		69000
				acagagtgag		69060
aaaacaaaca	aacaaaaaa	tattatttcc	aaagaaacaa	gaccctgggt	ccatttccca	69120
geecacacet	gatgttgatt	cacaacacac	ageetggttt	gctatgagcc	tgetteattt	69180
tatactasa	catatacata	tegetenne	agtcctggaa	taactctttg	ctgacctttg	69240 69300
accadacata	atctactac	accetcatce	ttaataatt	ctcactgcac aggctgactc	rgagicaara	69360
				ataaactcta		69420
				ctcgtgacag		69480
ctctttccca	ggatttcctc	ctctacctcc	tcaagtccca	ctgctctgca	aagaccaaaa	69540
				agcctcctct		69600
				cacccctctc		69660
				tgtagatgga		69720
				tgctggcacc		69780
				tagggatgtg		69840
			gatctttctg			69900
ctcctgtact	ggtgatccct	gtgttgattg	accactccct	tcctgggggt	cgtggtctct	69960
gtcccagttg	cccggacttc	tgtgagtgtc	ctactgaggt	ccttttcatg	agaagcatgc	70020
tgtccttcca	cctgctggga	gcaagagtga	caacttcaat	actataatag	cagtggcata	70080
				ccaagcagga		70140
aaaaacaaaa	acgtttacaa	gcaaactttt	tataaagggc	tagatagtaa	atattttagg	70200
ctttgagagc	cacatagact	tgtttgcagg	gactcaatgt	cgctattgta	gtttgaaagc	70260
				tcagcaaaat		70320
aaaacagaca	argagrgggc	tggatttggc	ccatgatect	tagtttgcca	actcctgctt	70380
rgggctcacc	cagatctgat	tetgaattet	ggctctgcta	ctggttagct	gcaggagett	70440
ggaaggetet	cigageetgt	ctcctcatct	gcaaaattaa	agcaataatt	tctaacactc	70500
taaaataaa	acttonogect	gradicccag	cactttggag	gctgaggcag	geggateace	70560
				aaccctgtct		70620
acacaaaaag	cagcogggda	ragragegeg	catttgtaat	cccagctact	rgggaggetg	70680

WO 01/27857

```
aggcagggat actgctagaa cctgggaggt ggagcgtgca gtgagtggag atcacacctc
                                                                                   70740
70800
ttagaaggtt ttgagataat gaataaaaga tgccttgtgt atactaagta ttcaacaact
                                                                                   70860
gatagetgea ttggtetaat tataacagtt tagaagegat tgagteaaca aatgetggat
                                                                                   70920
ttgtcaggga ggacttccta tcaggaggta gatcttgggc tgagtcctga agcaaagata
ggcattggat agaggagttg agagaacacc ctaggactgt tattattatt attcgacacg
                                                                                   70980
                                                                                   71040
                                                                                   71100
gagtetettg etetgteace caggetggag tgcagtggcg cgatetegge teactgcaac
ctctgcctcc caggttcaag cgattctcct gcctcctaag tagctgagac tacaggtgtg tgccaccaca cccggctaat ttttatattt ttagtagaga cagagtttca ccatgttggc
                                                                                   71160
                                                                                    71220
catgotggto togaactoot gacttoaggt gatocaccog cotcagooto coaaagtgot
                                                                                    71280
ggaataacag atgtgageca cegeacecag cecagaacea tttttcaate ettggetetg
                                                                                    71340
ccttttatta gctgcaagat ctcaggcaat ttatttaacc tctccaaaga ctcattttct
                                                                                    71400
cattcacaaa atgaggcaaa taataatatc tactatccca ggttgtcatg agaattaaat
                                                                                    71460
gcaacatgac atttaatgaa atgagaagtc ccttggacat taactggcta aagtatgtgc
                                                                                    71520
tcgacaagga tatcatttta ggtggatact tagcatctca gaactgatgc tcacaatgga atatcattga aacgcattaa aattcatttt aaatgattgt aggtagtgag gcaattgaaa
                                                                                    71580
                                                                                    71640
                                                                                    71700
gaagaagaca agaggactga ttataatgct tcaggctcac tagtctcctt ttaggaggga
aaaacaattt caagttaaat tttaggetet agatttttac ceetgetget cattagaate
                                                                                    71760
acccagattg atgaaatcag agcccatctg aggctgtgtt tttcatctcc agaatgagag
                                                                                    71820
ctgttgtggg gattaagtti tigaaaaagi acatciaaca ggtgatcgaa aatgatagtg
atattattgc agtgatggtc attattgttg ttattattat actgaaagag gcttcagttt
                                                                                    71880
                                                                                    71940
totgatocat aaagtgaggg aattgcatga gaccattgct aagattcctt ctagctctgt
                                                                                    72000
ttttttgttt ttgtttttta gacagagtet etgtegeeca ggetggagtg caatggeatg
                                                                                    72060
atettggete actgeaacet eegeeteeeg ggtteaaatg ateeteetgt eteageetee
                                                                                    72120
gaagtagetg ggaetaeagg cacacaceae catgeceage taaettttat atttttaata
                                                                                    72180
gaggtggggt ttcaccatat tggtcaggct ggtctcaaac tcctgacctc aggtgatcca
                                                                                    72240
cccgcctcgg cctcccaaca tgctgggatt acaggcatga gccactgtgc ccaaccctt
                                                                                    72300
ctagetttet tgateactga tietagggtt etetgetgaa atatatitga gacateetgg
                                                                                    72360
ataaaagatc atgcaagagc tcccaatatg gtattaataa ttgattctgg aggcttagct
                                                                                    72420
                                                                                    72480
actectgatg gattagaeat gacteaactg cetetettat gtgtacaaca caacaacaca
accaagaaag gttattctgg cattccattt attcagttta tttacagccc ttacttccag
                                                                                    72540
cagcacgtta aagatatggc cagggcggg tgcagtggct caagtctgta atcccaggac
tttgggaggc caaggtggc ggatcacaag gtcaggagtt tgagaatctg gcaattcttc
agacttagaa gcaaccagct cgataacaca gtcttgtgtg ggctctccct ctgtccctcc
                                                                                    72600
                                                                                    72660
                                                                                    72720
ctcgcttccc tcatttctca tccctgcccc tgagactgtg caccttcaca tagccctgcc
                                                                                    72780
                                                                                    72840
atgagacett cateteagge tttgetttet ggggtaactg aggetaaaca etgagtggee
ctaaaagagg attgggattt ggaagttaga ttattcacca gagaacagac tttgctgatg
                                                                                    72900
atcaggccca ggttgtaatt gttgaaaaaa agagaggatg catagtctta tctcatctcc
                                                                                    72960
                                                                                    73020
tagtcaaagt caacaccatg ataaataaga gtcaaatcct gagatgtgaa ttggggacat
                                                                                    73080
ttgagtggtt aaccetgaga agettgeace tteagacece teaatacece tgeteeceag
agaaggetgg acattgaeet cageacagge aggageeetg caagatgeea titgteetae
                                                                                    73140
taaagatgga cocctccact ctgtttctag gtaaataacc aaagtcaagt ctccacacag
                                                                                    73200
cctgagcaag aaagtcagag cctgctacag gagaaaatac cacactggcc aaaggattca ctagccctgg ccactgtgtg tgggaggaac cagggaatca tgtgtgggag tcaatgttga
                                                                                    73260
                                                                                    73320
agetgttgga etgggggtgg ggtggaatat aageetggee etggggagtt ttteeegttt
                                                                                    73380
                                                                                    73440
gagggeettt acceacaact caagatecag tgetatagea ggagatecca gagetagtee
taacagatgg tcaggattga acttggccta gagtaaaatg aggaggatag tgccagaact
                                                                                    73500
ttctcaacat actattgagg aagaggtcag aaggcttaag gaggtagtgt aactggaaag gggtcctgat ccagacccca ggagagggtt cttggacctt gcataagaaa gagttcgaga
                                                                                    73560
                                                                                    73620
                                                                                    73680
cgagtccacc cagtaaagtg aaagcaattt tattaaagaa gaaacagaaa aatggctact
ccatagagca gcgacatggg ctgcttaact gagtgttctt atgattattt cttgattcta tgctaaacaa agggtggatt atttgtgagg tttccaggaa aggggcaggg atttcccaga actgatggat cccccactt ttagaccata tagagtaact tcctgacgtt gccatggcgt
                                                                                    73740
                                                                                    73800
                                                                                    73860
                                                                                    73920
ttgtaaactg tcatggccct ggagggaatg tcttttagca tgttaatgta ttataatgtg
                                                                                    73980
tataatgagc agtgaggacg gccagaggtc gctttcatca ccatcttggt tttggtgggt
tttggccggc ttctttatca catcctgttt tatgagcagg gtctttatga cctataactt ctcctgccga cctcctatct cctcctgtga ctaagaatgc agcctagcag gtctcagcct
                                                                                    74040
                                                                                    74100
                                                                                    74160
cattttacca tggagtcgct ctgattccaa tgcctctgac agcaggaatg ttggaattga
attactatgc asgacetgag asgecattgg aggacacage etteattagg acaetggcat etgtgacagg etgggtggtg gtaattgtet gttggecagt gtggactgtg ggagatgeta
                                                                                    74220
                                                                                    74280
```

WO 01/27857 PCT/US00/28413

37/122

ctactgtaag atatgacaag gtttctcttc aaacaggctg atccgcttct tattctctaa 74340 74400 ttccaagtac cacccccgc ctttcttctc cttttccttc tttctgattt tactacatgc ccaqqcatqc tacggcccca gctcacattc ctttccttat ttaaaaatgg actggggctg 74460 ggcgcggtgg ctcatgcctg taatcccagc actttgggag gccgaggcgg gcggatcatg aggtcaggag atcgagacca tcctggctaa cacggtgaaa ccccgtctct actaaaaatg 74520 74580 caaaaacatt agccaggcgt ggttgcaggt gcctgcagtc ccagcggctc aggaggctga 74640 74700 74760 tagetgggca tggtggcgcg tgcctgtaat accagetact etggaggetg aggcaagaga 74820 atogottgaa cocagtaggo ggaagttgoa gtgagoogag atottgacao tgcactocag 74880 74940 cctggtgaca gagtgagact ctgtctcaaa aaaaaaaaa agaaaaaaa agacagaaag 75000 aaagagcaca gacagagtca caggtatttg cagtaggaag ctgtcaggtt agagtgcacg gaaatagaaa gtatatttta cacttacage acatettegt ttgattagee acatttaaaa 75060 tactgaatag caacgtgtgg ctatttagta ttcactaaaa tcttggacag tgcaagtcta 75120 75180 aagaateett gateegteeg geatggtgge teaegeettt aateeeagea etttgggagg ccaaggtgga aggatcactt aaggtcagga gttcgagacc agcctggcca acatggtgaa acctcgtctc tactaataat acaaaaaaaa ttagccgggc atggtggtgc atgcctgtaa 75240 75300 75360 75420 caaaaaaaa aaaaaattg ttgggcgtgg tggctcacgc ctgtaatccc agcactttgg gaggctgagg ggggtggatc acctgggttc tggagttcga gaccagcctg gccaacatgg 75480 75540 tgaaacccca tctctactaa aaatacaaaa attagctggg cgtggtggtg ggcacctgaa 75600 atctcagcta ctcaggaggc tgaggcagga gaatttcttg aacccaggag gcagaggttg cagtgagcca agatcgcgcc tctgcactcc atcctgggtg gcagagcaag actatgtctc 75660 75720 aaaaaaaaaa aaaaaaatac ttgattgtct ggacattctg cagaacatca tatggagaca 75780 ctatgttgac gacatcatgc tgattgtaag caagaaatgg caagtgttcc agaaacacag tcaagacaca tacatgccag aaggtgagat ataaactcta ctaagattca gtggcctgcc 75840 75900 acactggtga catttttaaa cctgctagat gtttgtgtag aaaaggattt aaccttgccc aaagagggt ctggcctttg tccccagcta ctggacataa tctctttaaa ctcttgaaat 75960 76020 76080 atcattcctg atagaagtat ttttgtittg actaggggcc ttggggccagc cagatagcaa caatgtgatc tgggttgggg gctttggatc aggtggcatc agtgtgacct cctgagtggc tagagactag aatcaaccac atgggcagac aacccagctt acatgatgga attccaataa 76140 76200 agactttgga cacaagggct tgggtaagct ttcctggttg gcaatgctct atactgggaa acceattctg actccatagg gagaggacaa ctggatattc tcatttggta cctccctggg ctttgcccta tgcattttc ccttgtctga ttattattat tattatgaga tggaatctcg 76260 76320 76380 ctctgtcacc caggetggag tgcagtggaa tgatctcaac tcactgcaac ctctgcctcc 76440 ccggitcaag cgaitticci gictcggcct cccgagtagc tgggactaca gatgcatace 76500 accacaccog gctaattttt ttgtattttt agtagagacg gggtttcacg ttagccagga tggtctcgat ctcctgacct catgttccgc ctgcctcggc ctctcaaagt gctaggaata 76560 76620 catgtgtgag ccaccgcgcc cagcccctt ggctgattat taaagtgtat ccttgagctg tagtaaatta taaccgtgaa tataacagct tttagtgagt tttgtgagca cttctagcaa 76680 76740 attatcaaac ctaaggatag ccttggggac ccctgaactt gcagttggtg tcagaaataa 76800 76860 gggtgctcat gtgtgtacca tgccctctaa ttttgtagtt aattaacttt cacaacttta ttattaccgc ttacactcaa tgtttattca catttatcca cataccactt attctagtgc 76920 76980 cttgcatcaa agactttcta tctcatgtac tttattctgc ttgaagtaaa tcctttagga tattettttt titttttaaa etttgeacat acataetttt atttttatt tattttaat 77040 tttgttattt ttgtgggtac gtagtagata tatgtattta tggagtacat gagatgtttt 77100 gatacaggca tgcaatgtga aataagcaca tcatggagaa tggggtatcc atcctctcaa gcaatttatc cttcaagtta caaacaatcc aattacactc tttaagttat tttaaaatgt 77160 77220 77280 acatttaatt ttgtattgac tagagtcact ctgttgtgct atcaaatata atttttttt tttttgagac agagtctcac tcagtggccc agactgaaag tgcagtggca caagctcggc 77340 teacticaat cicigectee etggiteaag egaateteet geeteageet eccaeatage 77400 tgggattaca ggcacacacc accatgccca gctaattttt atatttttt agtagagacg 77460 77520 ggttttegec atgttggeca ggetggtett gaacteetgg cetcaaatga tetgaccace tragectore anagtgetag gattaragge atgagerace acacetgger anaatagaat 77580 attetttagt gaggtetget ggtgacaatt titttettit tittgagaet gagteteget gttgteaget tgggetggag tgeaatagea egateteage teaetgeaac etceaeetee 77640 77700 77760 cggattccag caattctcct gcctcagcct cccaagtagc tgagagatta caggcaccca ccaccacacg cggctaattt ttgtattttt agtagaaatg ggggttcacc gtgttggcca ggctggtctc gaactcctga cctcaggtga tccacccacc ttggcctccc aaagtgctgg 77820 77880

WO 01/27857

```
gattacaage atgagecace aegeacagee aattttttee gtttttgtet gaaatettat
                                                                                                     77940
78000
tcattattat tattattttg agacagggtc tcactctgtt gcctatgctg gggtgtagta atgtgatctc ggttcactgc agacttgacc tcctagggct caggtgatct tcccacctca
                                                                                                     78060
                                                                                                     78120
geeteectag tagetgggae tacagatgea tgeeaccata cecaactaat ttttetattt
                                                                                                     78180
tttgtagaga tgaggctttg ccacatttcc caggctggtc tctaactcct gagctctagc
                                                                                                     78240
aatccaccca ccttggcctt acaaagtgct gggccatgac tagccagcag ttacttttta
                                                                                                     78300
tagcatattg aatalitaat atgaarcitc iggcatccac tgtaactgtt taaaaaatca
                                                                                                     78360
geligtttact tggcactett ttittttttt tiltttttga gacagagtet tgccctgteg
                                                                                                     78420
cccaggetgg agtgcagtgg cgtgatettg geteactgca agetetgeet eccgggttca egecattete etgeetcage etceggagta getgggacta aaggegeeeg ccaccaegee
                                                                                                     78480
                                                                                                     78540
                                                                                                     78600
cggctgattt ttttgtattt ttcgtagagt tggggtttca ccgtgttagc caggatggtc
                                                                                                     78660
togatotoct gacotogtga totgtoogoo toggootoco aaagtgotgg gattataggo
gtgagccacc gcgcccagcc totttttttt tttttttag acggagtett actotgtcat
                                                                                                     78720
ctaggetggt gtacagtgge gtgateteag eteagtgeaa cetecacete etgeeteage
                                                                                                     78780
ctgccaaata gctgggatta caggtgcgta ccatcacgcc cggctaattt ttgtattttc agtagagatg gggtttcacc atgttagaca ggctggtctc gaactcctgg cctcaagtga
                                                                                                     78840
                                                                                                     78900
totgoctgoc coagoctco aaagattaca ggoatgagoc accgcaccog gccaagtago actoottga aggtaatotg ottococtac coctagoaat tittaacaat tittottoat tittattico tgaagititg tiattaataa totgigigoa gatticititg tatticitit gittigoagit catagigata citgaattag tgigitiggit totgitatoa coacaggaaa
                                                                                                     78960
                                                                                                     79020
                                                                                                     79080
                                                                                                     79140
attgicagee gttagetttt caaatattte ettgetaaat tetetettet eccetttegg
                                                                                                     79200
tacaattgat ttgattaaaa ctaaaaccag ggccgggtgc agtgactcat gcctgtaatc
                                                                                                     79260
ccaacacttt gagaggctga ggcaggtgga tcacctaagc tcaggagttc aagaccagcc tggccaatat ggtgaaaccc cgtctctact aaaaatacaa aaattaccag gcatggtggc
                                                                                                     79320
                                                                                                     79380
acacattigt agtcaggagg cigaggcagg agaattgctt gaatccagga ggtggaggit
gcagtgagct gagatcccac cactgcagtc tggcctgggc gacagagtga gatgagaatc
                                                                                                     79440
                                                                                                     79500
tqtctcgaaa aaaaaagtta tgaatgtttg ataaactata tttgttagaa tgtttgttgt
                                                                                                     79560
agaatactat tcattgattt ttaaacaatg ttagattaaa ccattcactg gatttgtgat aattaactta ctgatttac ctcactgatt tgttgtaatt aatacaactg gtataaaaag
                                                                                                     79620
                                                                                                     79680
                                                                                                     79740
actgtgacga ggccgggcat ggtggctccc gcctataatc ccagcacttt gggaggctga
                                                                                                     79800
ggcaggcgga tcacctgagg tcaggagttc aagaccagcc tgaccaacat ggtgaaaccc
catctttact aaaaatacaa aattageegg tegtggtggt geatgeetgt aateecaget
                                                                                                     79860
cttcgggagg ctgtggcagg agaatcactt gaacccggga ggtggaggtt gcagtgagcc gatatcgcg cattgcactc cagcctgggc aacaagagcg aaactccgtc taaaaaaaaa
                                                                                                     79920
                                                                                                     79980
                                                                                                     80040
aaagaaaaaa aacacataaa acaaaacaac actgtgacgg ttcccaaaaa ttaggagcat
                                                                                                     80100
aattaaagga actcctgata aaaattaatt ttatcttaca tgtaaactaa aatgacttta
tgaagttaat tcagaaatac aatgcagggt attagtttgc cacagctgcg tattcagcct aatgtaatat tcttgttatt tttaaattct tcttttaact ttactcatat gtggatcatc aaatttcaaa agattaaatg acaatactct tagcagcaag cttccctaag catataaaca
                                                                                                     80160
                                                                                                     80220
                                                                                                     80280
                                                                                                     80340
ttttaatggg tgatgattca gaaggtaccc gaagaatatg tactgccaga tatcattcac
ccccatatac ctgcccgaca gacatcccat tttgggaccc tggataaatg tgtgggtgga
                                                                                                     80400
gagaaagata ggagaaagtg gtataagcaa atggctttgg agtctgattg acagcgattg aaatcctgtc tctacctctt aacagcctca tgatcctaca taagttaccc cgatcctcag ggccacatct gtaaattggg ggttgcgatg gcagccatct cacagggtct cttttcgggg aagggcagga attatggatt aagtgagcta gtaattgtaa agcacttaat acaaggaggg
                                                                                                     80460
                                                                                                     80520
                                                                                                     80580
                                                                                                     80640
cgcataataa gtacttcata aataatgacg gccattatca tgactgaggt gtatgcagct
gtcggggatt acggcgactt cagaatttct ggtgggcagg gctcaaaggc agcaaatcac
                                                                                                     80700
                                                                                                     80760
actggaagte gaggtgagge actgettetg cacagaetge ttagetggag agaatgagga aggettagag gagatttaga ggaacttaga gteeteegee teeaactetg tgggatetge
                                                                                                     80820
                                                                                                     80880
                                                                                                     80940
tecegtgeea gagacattea gaggatttet egeactetee ecteceetae gteceteeeg
coccatocaa ctaaccacac aacacataca aaatagcocc tgcgaggttc tgcacgctgg
                                                                                                     81000
aagggaacag gagaagggg ctgcgctttc ttgctgatgc cctgtacttg ggcccctggt agacacagcc acttgtcccc tcagcctgca gagaaatccc acgtagaccg cgcccgggtc
                                                                                                     81060
                                                                                                     81120
cttggcttca gccaatctcc ctttggtggg ggtgggatgc acgatccaag gttttattgg
ctacagacag cggggtgtgg tccgccaaga acacagattg gctcccgagg gcatctcgga
                                                                                                     81180
                                                                                                     81240
tecetggtgg ggegeegete agecteeegg tgeaggeeeg geegaggeea ggaggaageg geeagaeege gteeattegg egecagetea eteeggaegt eeggageete tgeeageget
                                                                                                     81300
                                                                                                     81360
getteegtee agtgegeetg gaegegetgt cettaactgg agaaaggett cacettgaaa tecaggette atceetagtt agegtgtgae ettgageagt tgaetttatt teteagtgee
                                                                                                     81420
                                                                                                     81480
```

WO 01/27857

```
81540
tagttttcca gataccagga ctgactccaa ggactattac tcatctggag ggtttagcac
agtaccgtcg catagtaaat ttccatgtca gttttggtta cctttcatgc acttgcaaac
                                                                                    81600
atgccatgct ctgaaacgaa ataggcacat ctttttttt ttttttta aggagtcttc
                                                                                    81660
ctctcgccca ggctggagtg cagtggcgcg atcttggctc actgcaacct ccacctcccg
                                                                                    81720
                                                                                    81780
tgttcgagat tctcctgcct cagcctcctg attagctggg actacaggca tgccacgacg
                                                                                    81840
cccaqttaat ttttgtattt ttagtagaga cggggtttcg ccatcttggc caggctggtc
taactcctga cctcaggtga tctgactgcc tcagcctctc aaagtgttgg gattacaggc
                                                                                    81900
ataagccact gcatctggcc agaaatgaaa taagtaaatc ttttaacctg ctctaacaat
                                                                                    81960
atagtgaaaa gaccatatta ttattagagc aggttaaggg atttgcctat ttcgggttct
                                                                                    82020
                                                                                    82080
agttatagtc ttaaacttgg acattettgt agaaagtaaa aagttteete tteaaagtte
cccttcttgt taaagaatac atcataagtg ttagaagtaa tagtttattt taaagactaa
                                                                                    82140
ctttcttcaa gcctccttgc tttgtgctaa taactctttg ttaagcccta tcctatgtaa ctgttggaca tgctcacagg cacgttccag ttcacagcct atgccccttc cttatttgga
                                                                                    82200
                                                                                    82260
aatgttattg cttccttaaa cctttcggta agcaacttcc tctccttctt cgttcttcct
                                                                                    82320
                                                                                    82380
tgcacttacc tatttagaaa gttttaggct attagcaaat cggctatcag tttaagagtg
tgaggtcccg ctccagccaa tggatgcagg acatagcagt gaggacgacc caaatgcgta agggataaat atgtttgctt ttcctttgtt caggtgtgct ctcgacatcg ttccatctgc
                                                                                    82440
                                                                                    82500
gattgagcac cctttctgca gaaagtaaag attgccttgc tggagatctt ttgtctccgt gctgactttt cttcgtggca ccgattatct atttctaaca attttggtat ttctaacatt
                                                                                    82560
                                                                                    82620
                                                                                    82680
ctgaacaatc ttgggctagt tgtctcttct gggcctgttt ccccatccgt cacatgataa
actteattgg tttaaaaace ceagegaaca tttattgagt taetattace tteetgeeet
                                                                                    82740
ccccaacccc aaccccaggg agcagttaca acctcagccg ctgagcgcac tegeegggtg
                                                                                    82800
ttaagaagca ccaaagacag ggaggcttga ttgattttgc tttgggagta gagggtcaga
                                                                                    82860
agattcacag gaaaatggca tttgagcaag gatgattcac tggagctagc ttttaaatac tggcgaggct tttatgttgc agtcccttac aaagttgagc attcgcaggg actgcactcc
                                                                                    82920
                                                                                    82980
gaaataagce egetteeet titeattege taatgateea gggagetget ggtteegeat geggeaggtt gtgeettte etaateaggg ttetgeateg eetegaacee geaggeegtg
                                                                                    83040
                                                                                     83100
gcgggttčte čtgaggaage agggaetggg gtgeagggtg aagetgeteg tgecggecag
egeetgtgag caaaacteaa aeggaggage aggaggggte gagetggage gtggeaggt
                                                                                     83160
                                                                                    83220
                                                                                    83280
tgaccctgcc ttttagaagg gcacaatttg aagggtaccc aggggccgga agccgggac
ctaaggeceg cecegtteea getgetggga gggeteeege cecagggagt tagttttgea gagaetgggt etgeageget ceaeeggggg eeggegacag aegeeacaaa aeagetgeag
                                                                                    83340
                                                                                    83400
gaacggtggc tcgctccagg cacccagggc ccgggaaaga ggcgcgggta gcacgcggg
                                                                                     83460
                                                                                     83520
gtcacgtggg cgatgeggge gtgegeeeet geaeeegegg gagggggatg gggaaaaggg
geggggegg cgcttgacet ceegtgaage ctagegeggg gaaggacegg aacteeggge
                                                                                     83580
gggcggcttg ttgataatat ggcggctgga gctgcctggg catcccgagg aggcggtggg
gcccactccc ggaagaaggg tcccttttcg cgctagtgca gcggcccctc tggacccgga
                                                                                     83640
                                                                                     83700
agteegggee ggttgetgaa tgaggggage egggeeetee eegegeeagt eeeceegeae
                                                                                     83760
                                                                                     83820
cctccgtccc gacccgggcc ccgccatgtc cttcttccgg cggaaaggta gctgaggggg
cgccggcggg gagtcaggcc gggcctcagg ggcggcggtg gggcaggtgg gcctgcgagg gctttcccca aggcggcagc aaggccttca gcgagcctcg acctcggcgc agatgcccc tgagtgcctt gctctgctcc gggactcttc tgggaggag aaggtggcct tcttgcgcga
                                                                                     83880
                                                                                     83940
                                                                                     84000
ggtcagagga gtattgtcgc gctggttcag aagcgattgc taaagcccat agaagttcct
                                                                                     84060
                                                                                     84120
gcctgtttgg ttaagaacag ttcttaggtg ggggttagtt tttttgtgtt tctttgagga
                                                                                     84180
ccgtggatca agatcaagga aatctcttta gaaccttatt atggaagtct gaagtttcca
aatgttgagg gttttatgtc taaaagcaac acgtgaaaaa attgttttct tcacccagtg ctgtcttcca atttcctctt tggggggagg ggtagttact gctgttacta aaataaaatt
                                                                                     84240
                                                                                     84300
acttattgct aaagttcccc aacaggaaga ccactacttt tgatgacttt ggcaagtttg
                                                                                     84360
                                                                                     84420
ctaactactg gaaccctaac ttacaaacga actacttaca tittigatti ccagtigtat
tacctgccca atgtttacgt agaaacagct taattttgat tctgggtaac gttgttgcac
                                                                                     84480
ttcattaaaa atacatatcc gaagtgagca agtatgggtc tgtggacagc agtgatittt
                                                                                     84540
                                                                                     84600
cctgtcaatt cctgttgctt cagataaaat gtaccagaca gaggccgggc gcggtggctc
                                                                                     84660
acgcctgtaa tcccagcact ttgggaggct tggcgggtgg atcacctgag atcgggagtt
                                                                                     84720
caaqaccaqc ctqaccaaca tqqaqaaacc ccgtgtctac taaaaataca aaattagcca
                                                                                     84780
gggtggtggc gcatgcctgt aatgccagct acttgggagg ctgaagcagg agaatcgctt
gaacctggga ggcggaggtt gcggtgagcc gagatagcac cattgcactc cagcctgggc
                                                                                     84840
                                                                                     84900
aaaaagagcg aaactccgtc tcaaaaaaaa agtaccagac agaaatgggt tttgtttct
                                                                                     84960
ttttttgttt tgagacggag tttcgctctt gttgcccagg ctcgagtgca atggcgcgat
ctcagtctcg gctcactgca acctctgtct cccaggttta atcgattctc ctgcctcagc
                                                                                     85020
ctcccaagta gctgggatta cccatgccc accatgcccg gctaattttt gtatttttag
                                                                                     85080
```

WO 01/27857 PCT/US00/28413

40/122

tagaaacggg gcttcaccat gttaggctgg tcttgaaccc ctgacctcaa gtgggcctcc 85140 85200 caccteggee teccaaagtg ceaggattac aggeatgage cacegeggee agecagaaat gggttttgga aaaagcacta aacaaaatcg aacttggttt catatgacag ctctgctgct 85260 aactgtaaca ggggcagacc agttaaccta cttttctgtc ttctgtcagc tgagaattag 85320 atgattccca aaggcccatt gaactctgaa tgactttaaa tacttcttct taagtgggta 85380 cacggttttg gtaactgatg ccaggtgatg aatgcatgaa agtgcttaat gaatgaaacc 85440 85500 ggtaaaatag taggaggaag ctttattggt aaggcagggg tatacctaat agctctctaa tttattggta ttgaagtggt taacttttgt ttttttaagg ggggaaaaca ttctaagaat aatgaggcaa actgcatatt gcacaagaga ctgttgtctc tattcaacaa ataccttttg 85560 85620 agtgtccaga gtctgccagg tgctgtgcta ggccctcacg attgagtagt gaaccagaga 85680 85740 atgiccetge acceatggag cttattgtet actggggtag acagataata aataagcaaa caaatettet etettetee tttegeteea tgtaagtgtg tgtgtatagg tgtataetta 85800 caagttgagt aaagtgttat gaaagattaa gaggagaaat gcattttggt tagatgttag aggactcagc aggtgacctt gaaacttaga gctgaaggat cagtaggagg taactagaga 85860 85920 85980 ggccagggaa tcgcatgttc aaaggccagg aggcaagaaa gagcatggtg cccttcaaga gaggaaagaa ggctactgtg actggagcat agatgtaggc aagtgttggg tgattgagag ctctacgggc catggttagg ttttattcct aatgccgaga tgccaaacat ggtggttcat 86040 86100 atctgtaatc ccagtatttt aggaggccga ggcaggaata tagcttgaac ccaggagttc aagaccagcc tgagcaacat gagacctgta caaaacattt aaaaaattgc tgggtatgat 86160 86220 86280 ggtgcacacc tgtggtccca gctactcagg aggctgaggc agaaggatca cttgagccta ggaggtggag gctacaatga gccatatttg agtcactaca ctccagcctg gatgacaaag 86340 tgagaccatg tgtcaaacaa aatacagaaa gaatattaat ttaaaattti gaaagaggag 86400 tgatctgaac ttatatctta aaaagatcat tctagggcat ggtggctcat gcctgtaatc aagggctttg ggaggctgag acaggaggat cacctgaggc cagttcgaga tcaacctgta 86460 86520 86580 cagcatagag agactccatc tctacaaaaa gaaaaaataa atagctgggt gttgtgagtt 86640 86700 ggtgcttttt ggaggctgga tgtggtaaga gtagaagctg gagatggtcc tgttagggat tcgattcaga ctttaaatac catcaatgca ttgagtccca aatttacatc actacgttgg 86760 86820 86880 atcettgeee etgaateeag actggtatat ceaactttag gtteagtttg tatetetace tgaccaatat agaggtgtcc agtcttttgg cttccctagg ccacattgga agaagaattg 86940 tottgagoca cacatagagt acactaacgo taacaatago agatgagota aaaaaaaato 87000 gcaaaactta taatgtttta agaaagttta cgaatttgtg ttgggcacat tcagagccat cctgggccgc gggatggaca agcttaatcc agtagatacc ttcaacttac aatatctaaa 87060 87120 87180 attitatioc agattiagte attitaaace toctcateag titticteaa gaagtagtat tttqqctttt tttcttttct tttttttgag atggagtttc gctcttatcg ttcaagctgg 87240 agtgcagtgg cggatettgg etcactgcaa ceteegeete etgggttcaa gtgattetee 87300 tgcctcagcc tcgcaagtag ctggaattac aggcatgcgc caccatgacc agctaatttt tggagacagg gtttcaccat gttggtcagg ctggttttgt actcctgacc tcaggtgatc 87360 87420 tgcctgcctc ggcctcccaa aggctgggat tacaggcatg agccaccgct cccggctgca 87480 87540 tittiggatt tittagttgct cagcccaaaa ctttagtaca tctttgaacc tcttctttcc toctactota tatotgatec atcagcaaat ctgttaggtc tacctcacac atatogaaat 87600 cctaccacgt ctcaccatct gtgacaatta acaccctggt ctaggcagtc atctctgtta 87660 agattgagig gttaaggatg icctctaagg agatgacait caaaictiag cttaaaigtc aagagggagc iggittiata aagattgagg aggcagcait attitgccai aggciiccai 87720 87780 ttggtttcca ttccattctt gatacttatg gtatatattc aaaacaaatg cacagaaaca 87840 gacccaggta tattgggaat ttcggatata gagttcctag ttgggaaaag atagactgat 87900 ctgtaaatga tgctagttat ccatcatctg gcaaaaaata atttcctgcc tcctctcata 87960 tatctcagat caacagactt tttctgttaa gggccaaatc ataaatattt taggctttcc 88020 88080 agaccatatg gtttctgtca cacteteett tateettgaa gecatagaca atatgtaaac aaatgggcat ggctgtgcta cgataaaact ttacttacaa aaactggtag tgggccagtt 88140 taggcatggc cagcactttg ggaggctaag gcagatggat cacttggggt caggagtttg 88200 agaccageet ggecaacatg gtgaaaceet gtetetaeta aaaatacaaa aaatagetgg 88260 88320 gcatggtggt gggtgtctat aattccagct actctggagg ctaagacaca agaatcactt gaacccagga ggcagaggtt gcagtgagct gagatagcac cactgcactc cagccagggt 88380 gacggagtct taaagcaaaa caaaacaaaa ggtagtgggt tgtatttggc ccatgggctg 88440 88500 tagtttgcca atccctgatg cagaaacaaa ttccaggtaa ataagagcct ggaatgttaa aaaaacaaaa cttgaagtca tgtagaagaa caggtagggg gaacaatcct gatctcagga taggaaggga tattgcttaa aataagacac aggaaaatat aatccatgtt gtgtaaattt 88560 88620 qactacgita aaacttaaaa ctttcgccaa gcgcggtggc tcacgcctgt aataccagta 88680

ctttgggagg	ccgaggtgag	cagatcacca	ggtcaggaga	ttgagaccat	cctggctaac	88740
acqqtqaaac	cccgtctcta	ctaaaaatac	aaaacattag	ccgggcgtgg	tggcgggcgc	88800
ctqtaqtccc	agctacttgg	gaggctgagg	caggagaatg	gcctgaaccc	gggaggcgaa	88860
				gggcgacaga		88920
gtctcaaaaa	aacaaaacaa	aacaaagcaa	aaaacctaaa	actttcatac	aataaagtat	88980
				gtgtatggaa		89040
				aaagaggaag		89100
caataaacat	gaaaggatgt	ttaacctcaa	aggaaacaag	gaaatgaatt	aaaaacatca	89160
aatgccattt	caaaactagt	aagttggcaa	aattaaaaat	accaaggatg	agaatatgaa	89220
				attaaaaatg		89280
				gctagaatgc		89340
				ttctcctgcc		89400
				aagttttgta		89460
				cctgacctca		89520
ttcccaaagt	gctgggatta	gaggcgtgag	ccaatoctcc	tggctgaaaa	aaatgcacat	89580
				agagaaattc		89640
				tttaagtgaa		89700
				atttataaaa		89760
				ctggttcaat		89820
				atttgtgtaa		89880
				ggagtgcagt		89940
				tectgeetea		90000
tagttgggat	tacaccacatc	caccaccata	cccacctcat	tttcctattt	ttagtgggg	90060
cagcegggae	ccatattaga	caccaccacg	ttgaactcct	gacctcaagt	attacacaca	90120
actteeete	ccatgctggt	aggetggte	acataeacce	ccgtgctcag	ccatttqcqt	90180
gatttttaaa	gatgtgcaga	ataataccat	taaaaaaaat	acacatacat	ctatatatat	90240
				agcactttgg		90300
				ggcgagatag		90360
				aggatcactt		90420
				caaagcaaga		90480
assastass	aataaaaaat	atttotatot	gatcatagta	aaaaaacgta	catogaagga	90540
aaaaaacaaa	atttatttat	ttattttt	ttttttaaca	cagagtettg	ctctgtcacc	90600
caggetgggg	tacactecat	taatctcacc	tcaccccaat	ctcggcctcc	caaattcaaa	90660
cagginggg	cacaguaguaguaguaguaguaguaguaguaguaguaguagua	totaectage	tagaectaca	ggtacccgcc	accacaccct	90720
cottation	gtettageet	tagagagaga	atttcaccat	gttggcaagg	ctactactac	90780
actectesee	ttaaataaa	cagagacagg	gacatacaa	agtcctggga	ttacacctct	90840
				tatttatatc		90900
agecacege	geteggeeag	aggaacteta	acctagedag	ataatactta	cadadacada	90960
agggtcaggt	ggtgattcat	aggaactcta	tattanatta	totatatott	2933340434	91020
				tgtatatgtt		91080
				ggtaaaagtg tgtggaaggt		91140
				gggtggatga		91200
				agaagatgac		91260
				agctatatga		91320
				tatctatggt		91380
				tgcagaagag		91440
						91500
				ccatgggcag		91560
				gtcgtcttcc		91620
				agtctccaca		91680
				ggagcatgtc		91740
				tctgagggcg		91800
				tacagagatg		91860
				ttaatgttga		91920
				tgatagaatt		91920
				gcttatctac		
				ttacagacaa		92040
atctttatt	ccttagacg	gagtetetet	trgtctccca	ggctggagta	cagegeegeg	92100
atctcggctc	actgcaacct	ccgcctcctg	ggttcaagca	attctctgcc	ccagcctccc	92160
				cattgttttg		92220
agagacgggg	LEECaccatg	ccggccaggc	tggtcttgag	ctactgacct	caggigated	92280
						_

WO 01/27857 PCT/US00/28413

```
accegeettg geateceaaa gtgetggaat tacaggegtg agecacegtg cetggeecag
                                                                                92340
acaaattatt atactctgag tgttagaggc ttaggatgtt ttcacttgat gctatgggag
                                                                                92400
gaataagtaa taagatatga tacacaacca aagacettte tteactatge ttetagtage
                                                                                92460
tagtactatg gatgacacat ggtaataata ttggttagca tttgtcctca atttactgtg
                                                                                92520
ctagttacte ttetaagece ettacaggta tatatttttt tteateaata ateetetaag
                                                                                92580
gtagttttta ttattgacct aattttataa atcaagaaaa ttaagaccca gagaagtaag
                                                                                92640
taacttgtcc aagatcacat ggcttataag tggtagagcc agaatttgac cccagatgtt
                                                                                92700
gtgactacat tgtctctcca taagcaggtt caactctttt gactggatgc tgttccaagg
tcacttcctt agagaagcct ttgctgacaa ctaccctcct gtgccctcct ccaaggctgt
                                                                                92760
                                                                                92820
                                                                                92880
ccattgttct agaactttga atactcatct tagaataaag ctggtctaat ttttacagtg
ttatagaatg gatctctgac tgcaaaagtt ggtcataatt atcttttat gttctagtga
                                                                                92940
aaggcaaaga acaagagaag acctcagatg tgaagtccat taaaggtaag ttctgccctt
                                                                                93000
ggcagtccac tgcattaaaa agtgatgtgc tttgcatttg tgagttcttt aatcctgtta
                                                                                93060
tactototot titggcatta atcatticig cottattita taattactta tgattitgat
                                                                                93120
                                                                                93180
ttatttccct ctttaacctg tataatgctt taacatctag catataataa gtaggctttt
ttttttttt ttttttgga gacggagtet tgetetgtta cecaggetgg agtgeagtgg
                                                                                93240
cgcgatettg geteactgea agetetgtet cocgggttea caccattete etgecteage
                                                                                93300
ctcccagca gctgggacta caggtgcacg gcgccacgcc tggctaattt tttgtatttt ttagtagaga cagagttca ccatgttagc cagtatggtc tcgatctcct gaccttgtga
                                                                                93360
                                                                                93420
                                                                                93480
teegecegee teggeeteee aaagtgetgg gattacaage gtgagecace geaceeggee
                                                                                93540
qtaaqtaqqc tttttttacc ttaattttat ttttttgaga tggagtcttg ctcttatccc
caggetggag tgcagtggtg ccatetegge teactgcage atccacetee egggttcaag
                                                                                93600
cgattetect geeteageet eccgagtage tgggattaca ggtggeegee accatgeeca
                                                                                93660
                                                                                93720
gctaattttt gtatttttag tagagacagg gtttcaccgt gttggccagg ccagtctcaa
                                                                                93780
actectgace teaagtgate cactegeett ggeeteecaa agteetggga ttacaggegt
gagecaccat geetggeeat aagtaggett ttactgagee ttgtgtgtat tggetateet
                                                                                93840
agtgattaca gtgaaccagt gcccttctta ttaatcacac atttaattgt tccctaaaag tgattagttc actttattta tttagtaaga caaaaaatga agaatactct taactgagca gtctgttaac tgtaggaaag cactgacact tataaggctt agttttctgt catttatcca
                                                                                93900
                                                                                93960
                                                                                94020
                                                                                94080
gaagtatggt tgattacagt ttttactttt ttatttgaat gaacaacctt aatttaaaat
ataltttgit tattttttgt tgggatcgat acattgicct igittataga ttagagcatg
                                                                                94140
ctttttaaag atgctgtatt actcactgat tttatttgtc cagtgtacag agattgaagt
                                                                                94200
gggaaaatta taatggaaat tgtttccata gtcattacat attaatttca tcaatttatt tccataaaat ctgtagattg ctacttattt agatttttcc ttcaaatgtt tttatgttgt
                                                                                94260
                                                                                94320
                                                                                94380
attgcttgca ctgagtattt attctatatg ctcaatttgc tggagaagaa gactaattat
aacttaggca agttgtaaaa ttagggaaaa aagtaaggta ccttacagcc tagtttactt
                                                                                94440
atttettatg taaageeagt tagatteeae attagtteaa actgeettet ttgageaaaa
                                                                                94500
cttgattggc agtgataaag gcttaaagcc cttctcaagc agagacctgt aaagactaga
                                                                                94560
                                                                                94620
tctgactgta gtagaaggaa ggaacttaga tgtttcaggc agtgagaaca ccagtcttcc
                                                                                94680
actitaaact ttgccactaa cagtatgacc ttgggaagtt gtaactttct tcagattctt
                                                                                94740
catttgttga atggggggat tggcctagct aatttctaaa tctctactgg gctaaaaaat
                                                                                94800
tetgtgetta tactetgatt atgaagtaca taatetgtge ttaacattea etgaettate
cttaggataa tacagaagca gtacaagaaa cagcccctca agatgtttgc agtctggtta
                                                                                94860
gaaagacaaa cttatacaca gaacagtagc aaatagacca aaataataat agctgccatt
                                                                                94920
                                                                                94980
tatagaacac ttcttctgtt ctgggcatta gacaaaaact gactataacg gtgaacaaaa
aagacttagg teetgeeete attgaactta cagattagta ggggagagga acattaatea
                                                                                95040
agtaettega cegategett agcetagatt ggtagtgatg gaagtaaaga gatgtgaacg
gacttgaaaa aaaattegga ggcaaaatgg atagaagttt attattgatt aaatatgagg
tgtgagagag agggatattt aagattgata eetaeettet ggettgeeta acagaaceaa
                                                                                95100
                                                                                95160
                                                                                95220
                                                                                95280
aacaggaaat tatatgttca gttttgttat gttgggtggg aggtgctttt gagtcattca
tttatatg ttatatgt tattttatat gcatagtaat tttaaggtct gagttttaaa ccaaaggtta gagagtgatt ttttagagtc tagcaaacct aagttgaaat cctgcctgtt
                                                                                95340
                                                                                95400
gaaatggctg tttactagct cattaaccta gggcaaagta ttcaacttgt tttcatttt
                                                                                95460
                                                                                95520
gtottoatot otaaaatgag gaaaatatgg tottacaaga ttgtootgag agatagatga
aataatatcc aaaaaaaaaa aaggtacata gagaaactcg talagtgcct ggtatatagt
                                                                                95580
aggiceteca tiggiageta teattateta gittitaaeat ageetteagi tigitgaatt agteaaaetg agigaageae tigeaaggaat teagaggaat tigagateaa eaaatgatit
                                                                                95640
                                                                                95700
                                                                                95760
ctgaagttta gggaagactt catggcaatg acacttacct tgtataaaag ttgaagaata
                                                                                95820
agaaagattt gaatgagaga ttetttetet tetecetace ageccagett ettatttgag
gatatattgg gcaaaggggc cttcagacaa gtagagggag atttttacag aaagattgag
                                                                                95880
```

atgaaggtat agaaggctgt aaagaccaga aaagagaatt gagacagagg aagcaggaag 95940 ccactgtagg titttgagca agatattgat gctgtaagta tggtgtttat gaaaggttag 96000 totggaagag atttgcagga tggagacccc ggaagttitt tigttataat acagaaagac 96060 96120 ttgcactgag ggtgaggtgt taaaaataaa caggtaagta aatgtttaaa catcttgaag gaaaagtcaa caaatcttgg caagtaaaca gataacagtg aaaaagaatg ggaccaagat 96180 tttgagtttt ggagactggt ggattgaaca gacagggaaa ttgagaggag aatcagatga 96240 tgatgtttta agttgatatt tagacagatt gtgcttgaga tggtaaagtc aatgtgggtg 96300 96360 ggaatgetta gtagegagta ateagtgata caagaceaaa geecaggtea aagacaagte 96420 acagatacag atcagggett titeatetge tecacagagg tgtaccetag gagetgttge aaacagtcca tgtggagggt gtgagtaaga tgtttccctt gaatttgcca gaattacttt tttgttgttg ttgttgttt ttctgagaca gattctcgct ctgttgccca ggctggaggg 96480 96540 cagtggcgag atcgcgcage teactgcaae etetgcetet egggttegag tgatteteet 96600 96660 geeteageet cecaagtage tgggattaca ggettgtgee accaageeca getaatttet tttgtatttt tagtagagat ggggtttcac catgttggcc agactggtct cgaactcctg 96720 gcctcgtgat ctgcctgcct cagcctccaa aagttctggg attacaggcg tgaaccactg cacccggtcc cttgttaagt ttattttggt gggaagcaaa ggaggtttca gcttttaaaa 96780 96840 96900 agtttgaaaa ttattgctct ggtaataatt aaagatttga gagtaaatat gctttctagc 96960 aqaaagaata aaagaagaac agatagcctc aagaagggga gccaaagaag caggctatat ctgacacact gggtgttgat aaatgggtat taaaagaatg agagcaatga gcagatagaa 97020 97080 gaggaaatta ggagagtata ataccatgga gaccaagaaa gatagactat caggaaggag 97140 tggtaaaaat aagttactag ttctaagaga gatgttaaga gggaccgggg aaagccttgt acaaatgagt tagtagcatt ttacattata tacatctaat taagaaacaa tgcgagagtc 97200 97260 teaceattee tatagactet taettgtact tgtetgaaca egaaaactgg ettttgttta taaataaget aaaaattatt ttgeteeaat tteteatgaa aataaaaata aacettettt 97320 taacattgaa aaaatagttt gaagacagtc actetteatt ttgtaattee cacaactatt attgaatgae tgaaattate tttattetga agecaaaggg gtgataetga tatttettea 97380 97440 97500 gactactaaa aatatattt atgaatttt agtgtgcttt atctttttt gttttttt ttgagatgga gtttcactce cgttgctcag gctggaggge agtggtgcaa tctcagctca ctgcaacett cgcctcccag attcaagcaa ttctcctgcc tcggtctccc aagtagctgg 97560 97620 gattacagge acetgecece acacceaget aattittigt attittagta gagacagggt ttcaccatgt tggtcagget ggtettgaac teetgacete aggtgateca eccacettgg 97680 97740 97800 cctcccaaag tactgcgatt gcaggcatga gccaccatgc ctggcctgag gaatattttt 97860 ctaggttccc cccacccaa gcatttattc tgcaatttta gttttgttcc taaagcaagc aaggtttaag gatttaaaaa taatccgtat titagaatgc tttctggctt tgttactttt 97920 tatecacagt agaagttete agagaatgat etecetett taatttaaet tittggeaca 97980 98040 gtattttgag aattataaat aatattagaa tgttttctgg ctgggtgtgg tggctcatgc 98100 ctgtaatcct ggctacttgg gaggctgagg caggagaatc acttgaacat gggaggcaga 98160 ggttqcagtg agccgaggtc atgccactgc actccagcct gggtgacaga gcaagactct gtctgggaaa aaaaaaaaa aaaaaaagag tgttttcttt cctattttcc accacttgat 98220 taagttactt ttcctcttaa gtattttttg ctgagtatgc tgacttaaga gtaatgttac aaaatttaat ttttaaagtt ctctgaaagc ccctttatga gagttttagg ctatcaaatt 98280 98340 gtgtttaatt cttaacaatt ttttgaaaaa ttatagcttc aatatccgta cattccccac 98400 aaaaaagcac taaaaatcat gccttgctgg aggctgcagg accaagtcat gttgcaatca atgccatttc tgccaacatg gactcctttt caagtagcag gacagccaca cttaagaagc 98460 98520 agccaagcca catggaggcc gctcattttg gtgacctggg taagtaacta tcattttta ttaacttgta ttagaaggat ttgagtacaa tatgtgaaac ttctgtcata ggatacagaa 98580 98640 98700 ctatataatt ggaaagtgct ttggaaaaaa tgtatttaaa ataacagcta caagtataat 98760 gggtagctgt gitgtgttcc tgtaaatata gaatataaag catgcccagt agaaaaacaa gcatttccag aagaaatata totgatcact aaatataaat atatgaaaaa gatgtotcac 98820 tttattactg agggaagtgc aaattaaaat aatcagttaa tgttctccta acacattagc 98880 98940 atatttttta aagtttgaca atttgaatgt cagtgaagat gcagggaaat acccctccta tttagtgata atataatetg gtgaagaete tttggaaage aatttggaaa teagtataaa atatgeatgt catttaggee actetteta agaeetagee eteagatatg eteatteata 99000 99060 99120 99180 tatgtatgta tgtatgttga aggctattca ttatagtatt gtttgtgata gcaaaaaatt 99240 atggacaaca tataaatate tgitataggg aaataaccaa attgtggtat acgcatgete 99300 togagtataa tatagccatt tgtttctatt tatttatttt cttgagacag ggttttactc tgttgcccag gctggagtgc agtggtatga tcatggttca ctgcagcctt cacctcctgg gcacaagcca ttctctcgcc tcagcctcca gagttactag gactgcaggc atgtgtcaccacacccagat aatttttaa ttttttgtag agacagggtc tcactatgtt gcctaagctg 99360 99420 99480

gtctcaaact	cctggcctca	agcaattctc	ccacacaggc	ctcccaaagt	gctgggatta	99540
ccaecgtgaa	ccaccacacc	tagttcagtg	tagccattta	gaaatctaaa	aaagacgtgg	99600
ganastatet	aaggatgtt	tasatatasa	aaaarraart	cacagtatgo	atootaaaat	99660
gaaaacgccc	aaggcatgtt	caaacgcgag	aaaagcaagc	accagoacge	tttattt	99720
ccgttatatt	aaaataagtt	cttccaaaac	aaaaacatat	gcaggagacc	Licatitige	
cagtatttct	tacccaaatt	tctgcactta	gaaaattgca	tgtcatgttg	tcataagttg	99780
aaaaaaagat	ccatgaacca	atggacttct	aataaaatca	gtcctgcttt	tgacatctct	99840
ctctactttt	gtgtatattc	aaaccagagt	gtcaatgtgt	ttgtggggca	cacttagcaa	99900
				aattgtaagt		99960
						100020
				gctctggaaa		100020
				actctgttgc		
gtacagtggc	gtgattacaa	ctcactgcag	cctcaccctc	ctcaggttca	ggtgatcctc	100140
cctcagtctc	ctgagtagct	gggactatag	gcatgcacca	ccacgcctgg	ctaatttttg	100200
tattttttt	ttttttttt	gtagagacgg	ggtttcgcca	tgtttcccag	gctggtctca	100260
aactcctqqa	atcaagcgat	ccacttqcqt	aggcctccca	aagtgctggg	attacgggcg	100320
				tgtttgtttg		100380
				ttttggctca		100440
						100500
				gagtagctgg		
				atatggggtt		100560
				ccgcctcggc		100620
gctgggatta	caggcatgag	ccgtcatccc	tggctggtgg	tttcttatga	cgtgaaacat	100680
gcaattacca	tatgacctag	cagttgcact	ctgtatttat	cccagataaa	tgaaaactta	100740
ccttccaata	aaaacctgtg	cacaaatgtt	catagcagct	taatattgaa	aaactggatg	100800
ttcttcacca	aataataa	ctaattcatt	cataccatoo	aataccattc	agcaataaaa	100860
				tcaagggaat		100920
						100980
				cgtttggata		
				agatgttaga		101040
tgagaggggt	aagtgggtgt	agttataaaa	gtgcaacatg	agggatcttt	gtgatgttga	101100
agttgtatct	tggcagtgga	tgcagaaatc	tcaatgtgat	aaaattacaa	agaactaaaa	101160
acaagaatga	gtatagataa	aactggggaa	atctgaacaa	gttagagtgt	tgtatcactg	101220
tcagtatctt	agagtgatat	tgtactatag	ctttqcaaqa	tgttaccatg	ggagaaacta	101280
				gatggggttt		101340
				tgccccagcc		101400
						101460
				gtattgtatc		101520
				aattgtaaac		
				ctagttgaca		101580
gactgaggat	gaagacacga	cgtgcttcaa	aaaaatgatt	tgaatatcaa	tggattaaga	101640
agaactcttt	tgacaaattg	atgaaaccct	cagtcagttt	tataagaatg	cccatcttta	101700
tgatcatgct	atgaaagcca	atttttaaaa	aaattttttg	tctttcctaa	caattagctt	101760
				ttttttatta		101820
				ctatgaatga		101880
				aggataagtg		101940
						102000
				gctatctgtt		
				cgaatgagag		102060
				aatctaattt		102120
aatgcatttt	atgttaattt	gcttttctgg	gatttttaat	gaggttgagt	atagttttta	102180
atattttat	tggccccttt	ggaactagta	tcataagttt	tttttcttaa	gaatttatgt	102240
agtctgggct	gggcgcagtg	gctcacgcct	gcaatcccag	cactttggga	ggccgaggtg	102300
				caacatggtg		102360
		tagctcagcg			cccagctact	102420
						102480
				ggaggttggt		102540
				gaaaagtctc		
				ttaaaagaga		102600
				tggacgattg		102660
agaaaccaaa	agcatagtat	ttgtagtttt	tttatttact	ttagttgcta	ggaagtaaac	102720
tttattcaag	gtctctggta	ccagttgttg	ctaaaagtga	ttgactaatc	tgtcaatctg	102780
				cttttaggca		102840
				accettgaac		102900
				cttcggcgaa		102960
						103020
				acttggtcgc		103020
acacageeea	aacacaguga	agcagagete	accggccgag	cctgtctctc	Califiadada	103000
						_

WO 01/27857

```
103140
gcatgaaact acagcgtctt ttttaactga ttctcttgat aagagattgg aggattctgg
ctcagcacag ttgtttatga ctcattcaga aggaattgac ctgaataata gaactaacag
                                                                                   103200
cactcagaat cacttgetge ttteccagga atgtgacagt geceattete teegtettga
                                                                                   103260
                                                                                   103320
aatggccaga gcaggaactc accaagtttc catggaaacc caagaatctt cctctacact
tacagtagcc agtagaaata gtcccgcttc tccactaaaa gaattgtcag gaaaactaat
                                                                                   103380
gaaaagtgag tatgtgattt tottgtgtgt acatatgtgt otcactttot ttttttaatt
                                                                                   103440
tactaagcag aacttcagat gaggaataaa atgattggaa tattttttt ctcctctaac
                                                                                   103500
                                                                                   103560
tacttgtaaa tttgggagaa tttggagagt gtagtagagt cagatcagtg tatggaaaag
gagcaggagt gactggacct tctaagaagt gtgttatcag aattagtaaa tgaagggtca
                                                                                   103620
aatgtoctae ttttoccoto cactgattit gacatcaaac cattatocac atagcottat
                                                                                   103680
tteeteete ggtettaatt ttattaatat tttaetgeae tttgeagata aaatttttaa
                                                                                   103740
aaaattttta aaaattgcca ataagtgaca tttattaagt tcagtgctta gtgtatattt ggattttatt tattagtcac aagacctttg tgcaggtagt aggcatgatt atcttttt
                                                                                   103800
                                                                                   103860
ttttgagatg gagtettget etgtegeeca ggetggagtg caatggegeg gteteggete actgeaacet eegggtteat gecattetee tgeeteagee teecaaatag etgggactae
                                                                                   103920
                                                                                   103980
aggegeetge caccacacce ggetaatttt tttgtatttt tagtagagae ggggttteae
                                                                                   104040
                                                                                   104100
catgitegee aggatggtet egateteetg actitgtgat eegeetgeet eggeeteeea
                                                                                   104160
aagtgctggg attacaggca tgagccaccg cgcccggact gattatctta tttacacatg
agaaaaccag ggcttagaaa ggttaggtaa cttcctctag gttgtacagt aaatgtggac ctagaagcat tttgacaaga gcacctgttt ttttttcttc tctattagtt tagaaattat
                                                                                   104220
                                                                                   104280
atactettaa ttateacetg ggattitgat tagacageet teatgitett titeatetta
                                                                                   104340
aatgttottt gtgtottaaa gggotaagtg atttottoag atottttagt toactoatto
                                                                                   104400
tcagtgaact aaaatgaggt ctaatctgct actgaatcaa gttttcagca tgttatttcc
                                                                                   104460
ttecteete cetecetet teetteete aaccaggete eggaggaget gggattacag
                                                                                   104520
gegecegeca ceaeteetgg ctaattttta tatttagta gagaeggggt tteaecatgt tggteagget gatettgaac teetgaeete aagtgaeeca eetgeetegg eeteceaaag
                                                                                   104580
                                                                                   104640
                                                                                   104700
tgctgggatt acaggcatga atcaccacac ctgacggcat gttattttca tcgcaaagtt
actgtaaget gggagaagig geacacactt gtacteecag etacteagga agettaaggt
                                                                                   104760
gagaagattg cttgagccca ggagttttga gaccaacctg ggcaacacag caagacccca
                                                                                   104820
gctcaaacaa agaaaaaaag ttattgaatt ttttatttct atggatcatt ttttgtagtt tcttattcct ttcacccttc attcccactt ttgatcccat cttttattta tttagttta
                                                                                   104880
                                                                                   104940
                                                                                   105000
ttaaatgtat atttgtctga taattctgct atctacagtt ttttgtggac ctgactcagc
atttctitgt ttcttcggat tcagactgtt ggtggcttgt gattttagtg atttttggcc
                                                                                   105060
gtgaacatgt ttcttggact tttgtctgtg ggaattetet gtgtactetg tataaattaa
                                                                                   105120
gttacticag gtgttttgca ttttcttttg ccatgcacct ggggcctggg tcactaccct tctggtacca cttaaaactg aatttttgtc ttgggtgctc gtactgatcc tgtatgagta
                                                                                   105180
                                                                                   105240
caggittata cttactgtag aaatatggtg titgattatg gggtattgtc ccagatggtg ctggagtatt aatatgctct ctgttaaact taatgtgttg tccctgtaaa actccaaaat
                                                                                   105300
                                                                                   105360
                                                                                   105420
totgaattoo agaatactao tggooccaaa tgtttaagat aagggcactg cotgtatttg
tttctgcctc ccactatttt ccttagttta acacaaactc acctttttaa aaaacatttt
                                                                                   105480
                                                                                   105540
gagagaatte agtattggga agagtiteta acetgtttet ggaaatggaa gtecaaagte
                                                                                   105600
tgtttctgta attgtttttt ttttgagatg gagtctcact ctgtcaccca ggctggagtg
caatgacgta ctctcagctc actgcaacct ccacctcccg ggttcaagcg attctcttgc
                                                                                   105660
ctcagecece tgagtagetg ggattaeagg tgeceaceae catgeetgge tgatttttgt attttagaa gagatggggt ttegecatgt tggeeagget ggtettgaae teetgaettt
                                                                                   105720
                                                                                   105780
                                                                                   105840
gtgatctgcc cacctcagcc tcccaaagtg ctaggattat gtttctgtaa ttgtaataca
tttattgttt ttagaaactg tctttgcttt agtggtaatt ttcaataaaa atagaaatag
                                                                                   105900
cagtggagtt attaaaagag cattagttac attittccct ttttcattat cttcaaatattatatatagt aagtttgacc tttttaaaat gtatacttgt atcagttta acacatacat
                                                                                   105960
                                                                                   106020
                                                                                   106080
agatteetgt aactgteace actataaggg taaagaacag ttagtteett cacetttgaa
                                                                                   106140
gtcaagecee acctetatee caacacttgg caacegetga tettteteeg tetcaatage
tttgcctttt ctctttttt ttcttatttt tttttttgag acagcgtctt gctctgtcgc
                                                                                   106200
ccgagetgga gtgeagtgag geaatetegg etcaetgeaa eetcegeete etgggtteaa geagttetee tgeettagee teectagtag etgggattat aggeaegeae eaccaeece
                                                                                   106260
                                                                                   106320
                                                                                   106380
ggctgatttt tttgtatttt tagtagaaat ggggtttcac catgttggcc aggctggtct
caaactettg acctcaagtg atccacctge ctcggcetee caaagtgetg ggattacagg egtgageeae tgtgeecaat caggaetttt tttttttaaa tttacattea acttgteatt
                                                                                   106440
                                                                                   106500
titticttgt atggattgtg cettcagagt cacacctaag agccetttge ctaagcaaag gtcatgaaga ttttctcata tgtttccttt taaaagtatt gtggttggcc aggtgccatg
                                                                                   106560
                                                                                   106620
gettatgeet gtaateteag caetttgaga agetgaggtg ggeagattae gaggteagga
                                                                                   106680
```

WO 01/27857

46/122

gatcgagacc atcctggcta atgcggtgaa accccatctc tactaaaaat acaaaaaaaa 106740 106800 aaaaaaatta geegggegtg gtggegggea eetgtagtee eagetaettg agaggttgag 106860 gcaggagaat agtgtgaacc cgggaggtgg agcttgcagt gagccgagat cgcgccactg 106920 ttttacactt tacgtttaga tatatatett ttttgagtta atgtegtata agtatgaggg 106980 ttacqtcaga ttttttgttt tttgtttatt tttacatatg gatgtctagt tgttctaata 107040 107100 ccatttqttq aaaagacaac ctttactcca ttgaattgcc tttgtacttt tgccatattt gtctaggcct gtttttggac teettttet gttteatgat gtgtgtet atteetttgt 107160 taataccaca tggtcttaat tactgtatag taagtcttaa aattgggtaa tgctggcctt 107220 ataaaacgaa ttgggaagtt tttattttta ctcttatttc cattttctag aagagattgt 107280 107340 gtagaatigg tgicattict tetttagata tttggttgaa ttgggaagtg atgecatetg ggcctagggt titgtttttt gtgtgtgaga cagagtetca ettetgteae eeaggttgga 107400 gtgcagtggt gagatettgg ettactgeaa cetetgeete ecaggiteaa gttateetee tgeeteagee teecaaatag etgggattae aagegtgtge caccatgeee gaetaatttt 107460 107520 107580 tgtattttta atgcagacag ggtttcacca tgttagccaa gctggtctcg aacttgtgac 107640 ctcaagtgat tagcccacct tggcctccca aagtgttagg attatagatg tgagccaccg tgcctggcag gggcctaggg tittcttttt cagagtattt taaactatga attcagatta 107700 tttaatagat ataggactat ttaagttatc tgtttcttct tgagtgaatt tttactgtag tttatggcct ttgagtaatt aattgtattg aattgtcaaa tttatgagcg tgtaattatt 107760 107820 107880 tatagcattt egggtttgta gtggtatece tettitatte etggtgttgg caattgtgte ttgtttttct ttgtcagatt gtatagggat ttattagtct tttcaaagaa ctagcttttg
ttttgatttt tctgttgttt tgttttcaat tttattgatt ttctgctctt tattattct 107940 108000 tttctattat ttctgcttgc tttgggttta ttttactctt ttttttttct ccaagttgct taaagtagaa acttagattt ctggtttgag acctttcttt tctaagataa gcatttaata 108060 108120 108180 ctgtaaattt ccttctaacc actgctttag ttacaccccc acaaattctg gtattttgaa ctgagcacaa atgaaatgtt ctaatttccc ttgaatctta ttcttttacc aatgaattat 108240 ttagaaatat gttatttagt ttgcaagcaa ttggagactt ttttcctgtt attttctac 108300 catttattte teattteatt atattatggt cagagaatat attttgaatg attteattta ttaattttta aaaataacat taaaaaattt tttaaaatgt gaatatacca catacagtat 108360 108420 108480 aaagattgta cattctgttt ttggacagtt ttctataaat gtcaagttga tttagttggt taatgatggt gttcagtttt totttattot tgctgatact ttgtatgcag ttatatcact 108540 ttattactca gaagagtgtt gaactttcca actacaattt ttttttccaa ttttactttc 108600 agctctatct ggttttgctt catgtatttt gaggctctgt tgttaggtgt gtacacattc aggatgatat cttctgggtg aattgcctgt tttatcatta tgtaattccc tctttatggt 108660 108720 108780 aalttteett gttetaagat cagaaatate tgttgteeaa tttatataga caetgeaget ttcatttgat tagtgcttgc atggcatatc tttttccatt tttttacttt tgatctacct 108840 ttataattot atttaaaggg ggcttottgt aggcagcata tagttgggta gtgttattta 108900 tttatttatt tatttattta tttatttatt tattgagaca gagttttgct cttgttgccc aagctggagt gcagtggtgc aatcctggct taccacaacc tccacctcct gggttgcagt 108960 109020 109080 gatteteetg ceteageete ceaagtaget gggattacag geacgegeae catgeetgge tgattetetg tattetagt agaaacggat teteaceatg teageeagge tegteetgaa 109140 etectgacet caggtgatec acetgetttg geeteccaaa gtgetgggat tacaggegtg agecaetgea eeeggetgag teatgttatt tttaatettt teteacaata eagggttttt gttggtaaat ttaattatt taatataat tttagtataa ttatttacat taaatgtaac 109200 109260 109320 109380 tgttgcactg gggtatttat aatgtgtaaa tataattatt ggtattaata taattatatt acticataata atattaatat cittggatti agattaccag titagtatat gitticigt 109440 ttctccctct ttgatttccc cttttttgct ttttttttt ttttaattct tatttttt 109500 tagtatttgt tgatcattct tgggtgtttc ttggagaggg ggatttggca gggtcatagg acaatagttg agggaaggtc agcagataaa catgtgaaca aggtctctgg ttttcctaga 109560 109620 109680 cagaggacco tgcggccttc tgcagtgttt gtgtccctgg gtacttgaga ttagggagtg gtgatgactc ttaacgagca tgctgccttc aagcatctgt ttaacaaagc acatcttgca ccacccttaa tccatttaac cctgagtggt aatagcacat gtttcagaga gcagggggtt 109740 109800 gggggtaagg ttatagatta acagcatccc aaggcagaag aatttttctt agtacagaac 109860 109920 aaaatqqaqt ctcccatgtc tacttctttc tacacagaca cagtaacaat ctgatctctc tttcttttcc ccacatttcc cccttttcta ttcgacaaaa ctgccatcgt catcatggcc 109980 cgttctcaat gagctgttgg gtacacctcc cagacggggt ggcagctggg cagaggggctcctcacttcc cagatgggc agccgggcag aggcgccccc cacctcccag acggggcagt 110040 110100 110160 ggccgggcgg aggcgcccc cacctccctc ccggatgggg cggctggccg ggcgggggct gacccccac eteceteceg gacgggegg etggceggge gggggetgac eccecacte ceteccagat ggggeggetg geegggegg ggetgecece cacetecete eeggacggg 110220 110280

						170240
cqqctgccgg	gctgaggggc	tcctcacttc	gcagaccggg	caactaccaa	gcggaggggc	110340
tecteactte	tcagacgggg	caaccaaaca	qaqacqctcc	tcacctccca	garggggrgg	110400
	gagacactcc	tracttoora	gacggggt cg	caaccaaaca	gaggcgctcc	110460
cggregggea	gagacactcc	ccagececca	9409999009	0350033300	3-55-5	110520
tcccatccca	gacggggcgg	cggggcagag	grggreecea	Cattleagat	gatgggttgt	
cqqqcaqaqa	cactcctcac	ttcctagacg	ggatggcagc	cgggaagagg	tgctcctcac	110580
ttcccagacg	gggcggccgg	tcagagggg	tcctcacatc	ccagacgatg	ggcggctagg	110640
	cctcacttcc	caascaaat	aacaaccaaa	cagaggetge	aatctcggca	110700
cagagacgct	CCCCaccccc	cggacggggc	9909900999	tagaggagta	agatgaggg	110760
ctttgggagg	ccaaggcagg	cggctgggaa	gcggaggccg	cagggagetg	agattacgtt	
actgcactcc	agectgggca	acattgagca	ttqaqtgagc	gagactccgt	ctgcaatctt	110820
gacacctcag	gaggccgagg	caggcagatc	actcqcqqtc	aggagctgga	gaccagcccg	110880
9900000099	cgaaaccccg	totocaccaa	aaaatgcaaa	aaccagtcag	atataacaac	110940
gccaacacag	cyaaaccccg			20225000	aggaaggttg	111000
gtgcgcctgc	aatcccaggc	actetgeagg	ccgaggcagg	agaaccagge	agggaggccg	
caqtqagccg	agatggcggc	agtacagtcc	agcctcggct	ttcacaactt	tggtggcatc	111060
agagggagac	caaaaaaaaa	gagaggaga	cqaqqqaqag	ccccttttt	getttettt	111120
ggattatttg	aatttttcct	taaatttatt	tatcttactt	atttatttat	ttttttgagt	111180
ggaccacccg		cacacacaca	casctacaca	catotoccac	tacacccadc	111240
gatteteetg	ccacagetee	Caaglagelg	ggaccgcagg	catgugueac	-t	111300
taatttttt	gtatttttag	tagagacagg	gtttcaccat	attggccagg	erggrerrga	
actettgace	tcaaqtqatc	cacctqcctc	ggcctcccaa	agtgctggga	ttacaggcgt	111360
gaggaggat	gccctgcctt	tttctagaat	ttatatattq	agttcttgat	tqtatctttt	111420
tatatagat	ttttagtggc	ttctctagga	attacaatat	acatactttt	cacagtgtac	111480
tatgtagget	Lectagegge	ttettetagga	actacaacac	2011200000	catagagaata	111540
tcacatttaa	tattttgtaa	cttcaagtgg	aatgtagaaa	acttaactac	Cataaaaata	
gaactaggga	tgaggttaaa	aaagagagag	aaaagaaatg	taataaagat	ttaataacac	111600
cattttttt	ttttttctc	tttttttt	gagacagagt	ctctcttct	gttaccaggc	111660
teeseteese	tggcgtgatc	ttaggtcagt	graaceteeg	cctcctgggt	tcaagtgttt	111720
Lggagtgcag	cggcgcgacc	ctggctcact	ttacacctcc	acacacast	cccactaa	111780
ctcctgcctc	agcctactga	gragerggga	LLacaggige	gegeeaceae	gcccagccaa	
tttttgtatt	tttagtagag	acggtttcac	tgtgttggcc	aggatggtct	cgatttcttg	111840
accttgtgat	teactetect	caqcctccca	aagtgctggg	attacaggcg	tgagccaccg	111900
cacccaacta	agtctttaaa	taffttttttq	acattgcact	ttttctcttt	tccttctagg	111960
cgcccggcca	cccaaatgtt	acttttctta	ttatttaaca	agttectasa	actttcctta	112020
attttagtaa	cccaaarger	agettegeta		tttatatta	totatatta	112080
cttctttaaa	tttttttc	ctgttgttca	gcttcgaaaa	tttttattta	Loughoutea	
aattcactgg	ttctttcccq	ttatttccat	tctqttattg	agtctttgta	gtgaatttta	112140
aattttgttt	attatgtttt	ttagttctaa	aattttcttt	ttttgtgtat	gtcttatact	112200
ttactactac	aactcttatt	tatttcagga	grgatchtat	ttettagage	atggttttag	112260
Legeteetga		teteeteege	Sestatet	cetetteatt	gtgttttgtg	112320
tagctactta	aaatttgttt	tateateeca	gcatatgtgt	cccccgacc	92222222	
ttgtgagata	atgggatttt	ctggttcttt	atatgacaat	taattttgga	ttgtatcttg	112380
gacagtttga	cttacqttac	atgattctga	atcttqttta	aatcctgtgg	aaaatattga	112440
agtttttgct	ttaacaagca	gttgacctag	ttaggttcag	tccacaaatt	ctaagcagca	112500
tteteteee	tctggttcca	tratractto	agttttgtat	cttatctqct	tatgtgcctt	112560
ccccgccggc	tetggeteta	teaccagece	terestees	angetttet	acacttttag	112620
tctgtgtcca	gtctgggacc	tggccaatgg	teaggreeca	aagcctttgt	acactttag	
aagcagggcc	atgcacaccc	agctcacgag	tggccccggg	agtgcacata	caactcgacg	112680
ttttcatggg	ctccttcttt	tctqtqatqt	ccctgacacg	ttctgccttc	taagaacctc	112740
cctttatccc	tttcctgttg	totagotaga	aagtcagggc	tttagattcc	ctatacttca	112800
	tgtagctatg	tecagetasa	+cccc3555	ttettettet	taggactaga	112860
geacacttee	igragerary	ccaacccccg	cygccacyac		teretere	112920
gtttctcttg	tcagaaagta	ggattcttgg	agetgetgte	attgetgetg	rageracter	
qatgctgcct	gggagtcgaa	ggagagaaag	gaacaaaaca	aaacaaccca	ggggatttcc	112980
tccactctct	ttgatccgtg	agageceet	ttcctgttcc	tcagaccaga	aatagagggc	113040
ctatattaaa	acttcttctt	tatacatcta	gratacagtt	tcagcttttg	agtccaggcc	113100
cegeeeegga		252500000	aggregation	gcaagttctg	aftacttttc	113160
aggaggtgct	ggacaaactt	gccaggagca	cggaggtact	gcaagtteeg	ttaattaa	113220
tcagtccacc	tgcttccaag	tecttggatg	catttgtcca	ttgttttgag	ttgcattcca	
taggagagac	agaagagtgt	gcttatttca	tcttgacata	cttattagga	tttcatatca	113280
aatcaaccca	tgatattctc	tatattaatt	tactattttc	cctttagcaa	gcacattagg	113340
	tttaacaccc	anatttaata	atttatata	taattattaa	tacttgactit	113400
addataatat	tttaacactt	gcccccggcg	geecegeea		costagaett	113460
tttttttt	tttgagacgg	agteteaete	tgtcctttga	ggcarryrcc	CCataaacct	
ttggtaaagc	atcaataatt	ttatctttca	tccacacaag	cttcaccata	aatttgatgt	113520
ttattcttcc	attttagcag	aattcatqtt	gctccaatag	gggctgtctt	caaactgatg	113580
ttttctcctt	cttagtgcct	cadadtadat	cctattcace	tacgttataa	caggttaata	113640
	ttastata	-agagtagat	nattentes	tagttttt	atcatatoca	113700
tgagtttatt	ttggtgtaaa	agractitga	aattcatgca	Lagittitt	accacacyca	
ttttccatag	ctttgaacac	ccccatgtaa	ctctcctctt	ccacaaacca	aacaatgaaa	113760
aagcaccttt	gtgatggaag	tttattttgc	aataggaact	cacagtgatc	taagccctgc	113820
tattcatgaa	tataattcat	tactggagtc	caagttgctt	tttgattttt	gaagttctct	113880
						_

WO 01/27857 PCT/US00/28413

48/122

tottocottg caggitataga acaagaigca gigaatacti tiaccaaata tatatotoca 113940 qatgctgcta aaccaatacc aattacagaa gcaatgagaa atgacatcat aggtaagcag 114000 tgcttgaaac tatggcaaaa aaaaaatgac aaaaaatgca cagaactgac aattttcgtt 114060 attgactaag ataatttttt ottaacatgg aatttagcag ttocottoot aatttgtttt otgagtattt tttatatogg attatagcto actttaaaag tttotoggot gcattoggtg 114180 cgagggtett tgcctgggcc agatgggctg cagtgtagcg ggtgctcagg cctgcccgct gctgagcagc cgggccggcg ggcggctacg ctaaccggca cagaccaccg gatggactgg 114240 114300 ccggcagcc cgcaccagtg cacgaagtgg gcgggacaga aacttctggg gttggaagtc cagtgaggct aaaagccggt accaaagtct ctaggcatca gggctgcagc ccaagagtct cacgaccagt gggcaactgg atggccagac aggtgtctca gtggtggcct ctccgtctca 114360 114420 114480 gggetteate ceaettetea gtgggeetga egteeetggg caeeetggat gtetaeetge attageeaga geeateaeat ggeetgtgae ttgeettttt ttgeeagttg attggeeae 114540 114600 acacagtgtc atttetgtgt catttggcac agetggaggt gcaaggagga gggcagcete 114660 atgtccagtc ccagtttcac gtaactttat tcttctgaat aaagacaatt tgctaacctt 114720 aaaaaaaaa aaaaaaaaaa agtttttctt atatgttgga cccaaattct taggctttaa 114780 cctgaataac aatgacagca agatcaataa atagtacaca tttattaaac actcactgtg 114840 toccagacaa tattocaago actittitatg gatagactoa tittaactio taaagaacti tgigggataa atacagitat titatagatg aagaaactga agcacagaga agitaagigo 114900 114960 tttgtccagg gtaacagctc agatatggca gagtcaggat ttgaaactag accetcacat 115020 accitaactg ctgtgctgtg gcagtgittt tcatactgta ggitgggacc agcettetet 115080 tatgcctca cccctgcca aaaaaaaaaa aaaaaaaaa aaatatatat atatatatat 115140 atatatatat atatatata aatatatata tatataaaat atatatatat ataaaatata 115200 tgtattagta tatatgcata tatagtatat attatatatt agtatatata ctaatatata 115260 atatacatat tagtgtgtgt atatatatat atactagaat aaaaaaatca aagtatctca 115320 gagtagtaag gacaaacatt tcagaaaaat gttttcatta tatatacatg tatgtatgtg 115380 tatgetgatt caacaaatat atttettata ggttatagea aaatagttig aaagetttta 115440 ctgtgtttta tcaggaagac cttaggtgaa cgtatattca cagataaaag aggttattta ttcattcaat aaatattaca ttctcataag tcctaatatt atgtatttt attcttcaaa 115560 aaagttagta tttgtgattt atgaaataag acatgttctt gcacttttag cagatctgtc 115620 ccgatgttgg gcttctttaa tccttagtgt gggtgctttg cactcactca ctgctgggga cagcaagacc cctgttagtc tcagctgtgt ttcttaaatt ggcccactgt accttccagt 115680 115740 tagetattet ggggteeatg teatgttgge tecattttee ttttettet cecacacaga tacetataae ggetataaca taggeetggt ggetgttggt ggettatece tatetgettg 115800 115860 tatttaaggg gtactgtttc actgagtttt gctgacagat gttgtcatga gatttgaggt 115920 tttctgtgtt gttgctctat ttttatgtgg gaalttgcta ctatcatcat ccctagacca gctttccta gtaatacaac agggatgttc tgactgatta gagtttgcct gtttgaagaa 115980 116040 ttggttggct agtgattitt ttttgagggg agtctgtacc agttaatagc ctgactggcg tgtggataaa aaggaagcag tttcaagtca aataaaacac ttaaaatgaa accacactgc 116100 116160 aactctcttt cttttactta agcttaatca aattaatgat gatgtaatcc catgaaggaa 116220 aagtettetg aaggateaag ttgataacat tttgtgatea aagaatttga gaaaacetet 116280 atcccagtgt ctatcattat atattttagg atgttaatta cctgtgtggc tttaggcaag 116340 tcatttttcc tccttgagcc ccattcttaa tcctgtccaa attatttgtc tcctcttgca 116400 gttggactat tttaatatag ctgtccttca agtgagtttt gttcaaagga gccttcactt 116460 116520 tagetettae tgtgtaceca etttgeatag tettgtttta aatgtaatee ttggattttt ggigttgcta actaattact gtttitatgi gaggatttag agtgatccag aaictatact 116580 tgcactacct cottoatott coacaaatgt ttgaagtggt agaattttta aaaactttga 116640 aggtacaget gacagaattt getgatggtt tggaagtgag tggtatgaga gggaaaaaaa ggaataaage atgactgeat tttttgtttg tttgtttgtt tgtttttgag aeggagtete 116700 116760 actetegeea ggetggagtg eagtggegtg atettggete aeggeaacet eegeeteetg 116820 ggttcaagcg attcccctgc ctcagcctcc caagtagctg ggactacagg cgctcgccac 116880 cacgeetgge taatttttt ttttgtattt tagtagaaac ggggtttcac cgtgttggce aggatggtet ceateteetg accteatgat etacteacet tggcetecca aagtgetgag 117000 gttacaggca tatatataag catataaagt gtgttatagc atacaaacag gtatatatat 117060 aaacatgcag tccacacagc tgataggaat gaggcagtag tgaaggagaa gttgatgtag 117120 gagaggggac agttgttaca ggaaagaagt ctggaggcag aagggatgaa ttccagtgct cacatagaag attgcttaga tgggagcaag gacaatttat ctagagtcac aggaaagaat 117180 117240 gcagtacacg ggtagagatg caggtgagtt gaaagatgtg agagatgatg gaaataattt tetgattget tetatattet caaggaagea ggaageaaag teeteageaa agagaataga 117300 117360 117420 agaggtgtta aatatttgag aaaggagatg tactqtaqaa aaaaaaaaaa ctcagtttct ccttetgaac teteacaaaa cagaaccett ecatgaetet agttgtgtgg ggttttttee 117480

				ccaactgggt	atcetetaaa	117540
ctgtcagcta	ccaattctgc	agatgattgt	tcagtgaaca	tacceggge	geeceeaag	117600
tcagttcagt	tctcacactg	tttacctgga	gatagcatca	gattettatag	accyaggacc	117660
ctataccaca	agactgcctc	cacttcagat	accaatctca	agtacaagtt	guggeergug	117720
arrataecta	accttctata	aattggagtt	cccacagtcc	cctccttqqq	ttcaataaat	
ttactagage	ageteteaga	actcagggaa	atoctttaca	tatatttacc	Callialial	117780
agaggatatt	acaaaggata	cagattgaac	aggcagatgg	aagagatgca	tgggcaaggt	117840
-+~~~~~~~~	ggcacagagc	ttccatgcac	tctccaqqtc	atqccaccct	ccaagaaccc	117900
ctacacattt	agctattcag	aagccccct	ccccattctq	tccttttggg	ttttttgtgg	117960
agacttcatt	atataggcat	gattgatcat	tagctattag	tgatcagctc	aaccttcagc	118020
agactcatc	caaaaaatta	gtgggtaggg	ctgaaagtcc	caaacqtqta	attetgeett	118080
	ctcattaccc	ctcatcctaa	agetettag	aggccacagc	cacaaqtcat	118140
ggtetttetg	gtgattaget ttgattaget	tccagagatt	ccatgaattt	raggegetet	atoctaagaa	118200
ctcattagec	Licadadadaa	aatgtctcag	ccatguatte	ccagcacttt	gggaggctga	118260
actggctaaa	ggccagctgc	aatgeettag	gcccgcaacc	taatcaacat	333433443CCC	118320
ggcaggagga	tegitteagg	ccatgagatc	adadccagcc	ttatatataa	teteagetae	118380
ccttacaaaa	aatttaaaaa	ttggccaggc	graaragere		atantagtag	118440
tcagaaggct	gaggatcact	gagccctgga	gttgaaggca	geagegagee	argaregege	118500
cactgactcc	ggcttgggtg	acaaagtgag	accttgtctc	agaagaaaaa	ggaaaaaaa	
aaaactgggc	aaagactaaa	taacatattt	cacagtatca	cagatttgta	ttgtctagga	118560
aagtgaatgt	aaacagacca	ggacactagt	atgatecett	ggtttcatga	aggtcccact	118620
aaagtcatga	acacaaagtg	agactaggca	tcatqttata	tggtttttcc	agccatgttt	118680
aacagctagc	taaatagcta	attotttcoc	tgcagtttat	tttagcagtt	ccttatttta	118740
gcacatttca	tottttaaaa	tttctaccaa	taacatttta	ataaactttt	ttacagataa	118800
cttcacaaat	ccataatttt	ttaagttaca	atcccagaaa	tagaattgct	cattgaaagg	118860
gratgiteat	ttttaaagtt	atoctagaaa	ctqccaaatt	gccttcagaa	aaaggtgttt	118920
gtatgcccac	taacactagt	gttagttttc	ttataccctt	gctcaagtat	acatattatt	118980
aaaacaato	rragaccagt	ttactagata	aaaggtgtag	tgcctcctta	ttctaatcta	119040
tttcattact	agraagratg	tatgtctttt	cacattaatc	attttatqtt	tgttcctttg	119100
teasttates	tatactttac	tcatttttct	trtggaacat.	ttcttagtag	tttataaqaq	119160
ctattagtea	tttaatgata	gtaacctttt	aactgtcatg	catoctocaa	atctttttc	119220
tetteggtat	cctttctatt	ttgtttttgg	agggtttcta	totataggaa	ttaaatttta	119280
tgtttgtttg	tottttgtatt	tctgcttttg	catatotact	tcaaaagact	ttctatttta	119340
Egitgitaaa	theatetat	tttcttttag	ttctatttaa	aacctcttaa	tttatatgcc	119400
agatcaagtg	ctacctgtat	tgattcacaa	atatatatac	atactttcaa	tttagtggca	119460
tgtgctgtta	acteccaagt	ttttgcagca	aggegeacae	cacacacata	acaddtagat	119520
atttaattat	ttacaacttc	ttttgeagea	aggatttgtg	gagaagacgg	acaggeggat	119580
cccaactgtt	tegttttgge	acagtccata	geoccageg	caatygagca	agagcaagcc	119640
agttcatatt	ttcacattgt	gcatcctagg	gaarrigggi	ttaccettag	tacaagatca	119700
cactcagcta	aaaacaaagt	atttttgaga	atttaaatat	LLLggalact	cacaagacca	119760
tataaagcat	actctatctt	ggttaacagt	ttcttttaaa	tataaattat	grgaacter	119820
aaaattttca	ttttcatttt	caatgttaat	atttcctaag	ttaaaataat	tegetettag	119820
ttctgaaata	atttggggag	tgattgagtc	tgtagtgatt	atgactatta	gaactggttt	
atttatttaa	ataatgcatg	tcttcagatg	gctctcctaa	tttgttagtt	aggetttaag	119940
ctasatggat	octatataac	taaatccaca	tagatttgtt	gaaatggctc	cagaggtttt	120000
ttagatttat	tactgctatg	tqcccttaaa	aaaaatctat	tcattcttc	acttaacatt	120060
tatcagaaga	atactetata	taagacgtgg	ttaqqcataq	tgccagtctt	gaaggaagtt	120120
acageetaat	aaaagacata	gggcatgttg	tttggttact	gtaatatgaa	gtggcatgtg	120180
ttaaatgtca	ggggagaact	acaaagtcat	aaaaaqqtqq	gagagattac	atacaggtaa	120240
aggaatgagg	aatgacacca	·tagagagtaa	aataatatta	acctaggeet	ttaaqataca	120300
atagggacag	tatggaaaga	gtatatttt	cccacttaaa	ctctttcctt	ggtcgttccc	120360
tcaaattttc	ccttttatcc	atgtgcaggc	actttagtga	gtttctgcga	agtcaccatt	120420
tototaaata	ccadattgaa	gtgctgacca	atagaactat	ttacctqqct	gacattctct	120480
tototoacto	acceptett	tatttctctg	aggtaaagtc	tocatttctt	ttcacactct	120540
attagagagag	tecaectet	aactatcaat	actagaaccc	totctatago	aaataacaca	120600
accogagoat	ctcaycccc	aaaagatgta	tcattotttc	aagttgtttc	tgatggcaag	120660
gaagagccaa	greattee	gagagaacat	caaaattcaa	totattaaat	aactctaatt	120720
agtaatttaa	Laatalatta	gagagaadat	gaaaaccaa	totttttaat	tatttggagg	120780
ttgagaaacc	caaccaaacc	actgcatgta	ayayaytyca	tanttant	accacactto	120840
tattttaaaa	ccacagaatt	tgaaacttgc	trecagigea	caaactycay	tagtgacttc	120900
agaagagaaa	aaaagtagta	aattttttct	tatgctcatc	acciciaci	tatttta	120960
ataggattgc	ccagtgaaga	agcatttgca	acagacaatg	agtatattaa	tctttttgag	121020
gcatacagtt	tagtataatg	ctctttgtta	ggcttcaaca	agtgaaatta	ttttgttgga	121020
aaqcaaatqa	ctattaagta	gaaagaggat	tcccagtctc	acaaagcagt	aatttagaca	121000

WO 01/27857 PCT/US00/28413

50/122

ctcgattctg cctctttaca agaatacagg tactcagttg atttgttttc tcactccctt 121140 totttgctat aagtttaaat caacaatttg tttaggttaa tatgtcctca tggaatggtg 121200 gaaatgatca gatataaaat atttggtttg gttagtttac tetttatatg titgetggca 121260 121320 aggaaccaca aatccagttt agtataattt ttactctagt tcactaaaag tttgcatcca gctgtgtagg tagtgtttgt ttcttgttaa ctttttttc gtctaaaaga atactttaaa 121380 actiticaat cicaaatgac tgtaactigc tgacaggtgt taacagaaga agtagatett 121440 tttgtttttt gcttatgacc tgtattttaa tatttgagct tatagattag agattgtgag 121500 agaaatetgt ttatagtett atttteeett gtgtattttt tetteetagt acatggaaaa 121560 agaggatgca gtgaatatct tacaattctg gttggcagca gataacttcc agtctcagct 121620 tgctgccaaa aagggccaat atgatggaca ggaggcacag aatgatgcca tgatttata tgacaagtga gttatattga tagatggatt cagcagatac ttattgaaca tttgatatgt 121680 121740 tttgtggaaa taaagatgaa taaactcagt ctctgttgtc aaggagctca caggaggcag 121800 cataaaagct gcttttatat ggtgtttgta aagctttggg ggttcttaga acaaaagttt 121860 ctgctgggaa aggggaggtg tatgtggggt aaacaggatg gcaatggtgg tgttcaagga 121920 gtgtttccca gaagagagat tttgtttgga tcccaaagaa agaagggaat tttgctaccc 121980 agagaaggca gaaaacaaca ttctaggcaa aggcattggc ccagaagcca tggaaacgta 122040 122100 ggggaaagtg gcactttcaa gaaacttgag tttagataat caaaggagtg gggaataaat atgaggatgc tggtactaat tggaatagat tgtaagggac cttgaatgcc tatttatggg 122160 tatattatac titictgtata aatctgctca ggcacgttgt taattagttt tttattagtt 122220 ttcactgaaa atgagaggat ggaaacatca tacagtaaac aaaattgaaa atatctggtc aggcagatga tgagcttgtg gccagctctg taacgtatgg tattcttttc atttaacttt 122280 122340 122400 tettactetg taaaaaaagt aattegtggt egggeaeggt ggeteaetee tgtaateaca acactttgag aggcagagge aggtgaatcg cttgagccca ggaatttgag accagcctgg gcaacatggc aaaacccgcc tttactaaaa atacaaaaat tagctgagcg tgatggcgtg 122460 122520 cgcctgttgt cctagctact taggggcctg aggcagaagg atcacctgag ccttgggagg 122580 122640 tegaggetge agtgagetgt gatecaetgt actecaecet gggeagggea gtagagtgag accetgeete caaaaaaaaa aaaaacaaca aaggtaattt gttatttgta teettaagca 122700 aatgctaaag gggtaacttg gggatagaga aaagtccaca gatgttaggg tttgaagacactaatagtat ctaggccagt ggttcctgaa cattagtctg tgggctcttg ctgggctgtc 122760 122820 tgcataggaa tcacctgaga gcttattaaa aataggtttt caggctggtt gcggtggctc acgcctataa tcccagcact ttgggaggct gaggcaggcg gattacttga ggtcaggcgt 122880 122940 123000 tcaagaccag cctggccaac atggtaaaac cccgtctcta ctaaaaatac aagaattagc caggeatgat ggeacacace tgtaateeca getaeteagg aggetgagga aggagaattg 123060 ctcgagcccg ggaggtggag gttgcagtga gcggagatca tgccactgca ctccaggctg 123120 123180 coggtggctc acacctgtaa toccagcact ttgggaggcc aaggcaggca gatcacttga 123240 ggtcaggagt ttgagaactg cctggccaac atagtgaaac cttgtctcta ctagaaacta 123300 caaaaaatta actgggcatt ttgacgggtg cctataatcc cagctactag ggaggctgag 123360 gcaggagaat tgcttgaacc cgggaggcag aggactgcat ctcaaaaaaa aaaaaaaaa aaaggtttoc agtccccctg tctcagaaat tctgattctg caggtttgag gtgtgaccag 123420 123480 123540 gaatetttat ttttagaaga cataccagat aattetgata aatagecagt ttagggatgt agtotaattt tootattttg caagtaagga aaataaggoo cagagaggta atgattttot 123600 caaagtcaca gaacaagtta gtggcagaat ttggactgga atgcagttct taatgttctg 123660 tccagtgttt attctggtac agtatgtttg tagaaggtat tacgtaagaa acattgttat 123720 atagatgttg agataggaag agtttacatt tagaaatttg gtctaaaatg cctgaacatt caagtcgtgg aggagtattg accaacttac tcaatacaac ataggagatt cacattttgt 123780 123840 123900 tacaaaaatg ctgatttaaa aggagagttt tetttttttt ettettttt atttttgag atggagtett getetgteae ceaggetaga gtgeagtgae acgateteag eteaetgeaa eeteeacete etgggtteaa geggttetee tgeeteagee teetgagtag etgggattae 123960 124020 124080 aggtgggggc caccacgccc agctaatttt tgtattttta gtagagacag ggtttcacca 124140 tgttggccag geeggtettg aacteetgae etcaagtgat ceacceacea etgeeteeca aagtgetggg attataggeg tgagecactg tgeecageet gettgttttt gtateatata 124200 tatgcatcat cataatcatg cattatcaac ctttgtattt ctgtcaggac atagaaacca 124260 ttagagtgct tggaagagag ccttttttt tttctcgcat ttaatgcttt ttttggtatt 124320 124380 catttcataa tcagcttacc aaaacattac ctgcattata ccccatcaag gtagaaatct ttgtgttatc aatattggtt actocottto cacacogagt catcagtaag tootgttota 124440 tocaaatagg toatatgcat otagetoace ceteagiget gttttgttti gaatitgtac 124500 124560 124620 124680 ctatctatca agtoctcaac agagaatagg tacccataaa tatgtgattg ttagtttctt

				-at agast to	2000102202	124740
tgcctcagtt	gtagtctgat	ccttacagct	tttaaacaac	agtagagttc	teteteraga	124800
actaaggatg	gttggcaggc	agatagaaag	gtagcaagtt	gacccaacta	-ttetagggga	124860
antoqqaaca	aagaaaggtt	acatcagcac	totcatcaca	Lagetetata	gillaggel	124920
tgcaggctca	atcaagtagc	cttgtataag	attctctgga	ggaggtgctg	aaagttgctt	124980
atacttocta	togaatttga	ttttacttcg	gatatettt	taccataggt	actteteect	
0000000000	catcctcttq	gatttgatga	tattatacaa	ttaqaaattg	aatccaatat	125040
ctacacacaa	ggt.gggccac	tccccaactq	tttcacaact	ccattacgtc	aggcctggac	125100
aaccatggag	aaggtaaccc	agaacttcaa	acqtatcaaa	ctacaagaag	ttttattggt	125160
acaactcata	aaatataagg	toggaaaacc	aaqcaqaata	qcacaqtqqa	aattgaagca	125220
atccaacaaa	gtgattaaga	acagagact	tgagtctggc	ctggtatgta	cagtcacgtg	125280
ccacataaca	ttttagtcaa	cagtogacto	catatacaat	gqtcctqtac	gattataaty	125340
ccacacaaca	ccttagtccaa	taataacaaa	agttagaaaa	aataaatttt	aataaqtaaa	125400
gattaaagtt	ggcagcgcaa	aaacataaaa	gaataaccaa	gaacaaaaca	aaaaaaatta	125460
aaayaaaaa	gaaaaaccaa	tattacata	tatttactot	actatacttt	taatcattat	125520
taatggaget	gaaaaacccc	tactaacaaa	acacttaact	gtaaaacagc	ttcagacagg	125580
tttagagtgc	CCCCCCCCCC	caccaagaaa	attatcaaac	gedatasego	ctccatgcgt	125640
tccttcagga	ggtttccaya	aggaggcacc	greatetee	gagatgacgg	aagtgttatt	125700
gttactgccc	ctgaagacct	cccaguggga	caagacgcgg	aggtgaaaga	tttttaacaa	125760
gatgatcctg	accetgtgta	ggcttaggct	aacgcgggcg	tttgtcttag	taaggatata	125820
acaaatttaa	aaagaaaaaa	aaaaccaaaa	atagaaaaa	gcttataaaa	aggatata	125880
atgaaaatat	ttttgtacag	ctgtatatgt	EEgegeeeta	agctgttatg	acaacagage	125940
caaaaagcta	aaaaaagtaa	aacagttaaa	aagttacagt	aagctaattt	accaccaaag	126000
aaaaaaattt	taaataaatt	tagtgtagcc	taagtgtaca	gtgtaagtct	acagtagtgt	126060
acaataatgt	actaggeett	cacattcact	taccactcac	tegetgaete	acccagagca	
acttccagtc	ttgcaagctc	cattcatggt	aagtgcccta	tacagatgta	ccatttttta	126120
tottttatac	totattttta	ctataccttt	tctgtatttg	tgtttaaata	cacaaattct	126180
taccattgca	atagtggcct	acgatattca	ttatagtaac	atgtgataca	ggtttgtagc	126240
ccaaaagcaa	taggttgtac	catataqcca	aggggtgtag	taggccatac	catctaggtt	126300
totataaqta	cactctgtga	tottaocaca	atggcaagca	gcctaacgga	aattctgttt	126360
attgattgat	tgattgattg	attgattgag	acagagtttc	actccattgt	ccaggctgga	126420
atacaattac	acagtettgg	cacactgcaa	cttctqcctc	ccaggttcaa	ccaattatcc	126480
tacctcatcc	teccaagtag	ctoogattac	aggcaggcac	caccatacct	ggctaatttt	126540
tatatttaa	tagagacagg	gtttcaccat	tttggccagg	ctqttctcqa	actectgace	126600
ttaagtgatc	tacctacttt	ggcctccgaa	aqtqctqqqa	ttacaggcat	gageractar	126660
acctagacag	taactgaaat	tctctaatqc	cattttcctt	atctgtaaag	tgacgataat	126720
atacacattt	acctcaaagt	tactttgatg	attaaaqtaa	ggtaatgtat	ataaaataca	126780
tattaacata	gtacctgaca	catootaaoc	atcaaaaaat	gttaactact	tttattacta	126840
ttattattac	gtattttaa	ataattagag	agcagtatca	aaaattagct	gggcgtagtg	126900
ccatcatcat	atanttocan	ctactcagga	ggctgaagct	ggaggattgc	atgageetgg	126960
gcatgcatct	ctagetteag	ccatattcat	gcccctgcac	tccagccttg	gtgacagagc	127020
gaactaaagg	cttcaacaat	tasagaagg	attatgccgc	aacgttagct	tagaaatgat	127080
aagaccccgc	coccactasc	tatagaagge	attogaacco	tagttttggg	tattatgatc	127140
ccacacacac	attantage	tattaataat	aaaacattaa	ctaggcacgg	cgactcacat	127200
acaaggtatt	accaacagee	caccaccac	taagtgata	acctgaggtc	aggagtttga	127260
ctgtaateee	ageaetttgg	gaggeegagg	tototactac	aaatacaaaa	ttagccgggc	127320
gaccagcctg	accaacatgg	agaaacccca	ttaggagge	gaggcaggaa	aatctcttga	127380
gtggtggtgc	atgeetgtaa	teecagetae	ccaggaggcc	ttacactaca	acctagacaa	127440
acccgggagg	cagaggttgc	agegageega	gattgcatta	ttgcactcca	taaagtatta	127500
caagagcaaa	actccgtctc	aaaaatataa	ccacaacaaa	taaataaaag	caaagtattg	127560
atgtttgtga	atgatttatt	cttctaatga	actagaggag	atttttccag	acttaattaa	127620
gccagtgagg	ttatgttgct	tgtatgtgtc	atgtgtatee	aggtgaaaaa	acctaactaa	127680
acgctattat	ataataccat	acataaaaac	tgaattttag	gaatactgaa	gaatgacata	127740
tagaagtcaa	atcattaaat	agctagtagt	aaacagaata	gagtgtcagc	tgttacccaa	127740
tgatgataat	attttcacga	ttaaaattaa	accttttctg	attttaaagg	aaaagttcag	
atctgtatca	tataaaqaat	gtaaattttc	agggtaataa	aattaaaatg	cagagagaaa	127860
aatgcaaaaa	tagttcttac	tagatgtgtg	tatqtaaqqa	acttagacta	attttaagaa	127920
cactgtcaag	accetagtag	ttaggtagga	aaaaagacat	gaatgattca	ttcaacaaaa	127980
actttgagta	tttctatact	agatggtagt	gttacagtgg	taaacaaaat	aaatgtgttt	128040
ctoctatcct	ggagcttagt	ctacaaaaaa	ggtacatatt	ggccgggcac	ggtggctcac	128100
geetgtaate	ctagcacttt	qqaaqatcga	ggcgggtgga	tcacctgagg	tcaggagttc	128160
aagaccagct	tggccaacat	qqcqaaaccc	cgtctctact	aaaaatacaa	aaattaactg	128220
aatataataa	cggacacctg	taatcccagc	tactcgggaq	gctgaggcag	gagaatcact	128280
JJ - J - J J - J J		-				

WO 01/27857

tgaacctggg	agacagaggt	tccagtgagt	cgagatcatg	ccactgcatt	ccagcccggg	128340
ggacaaaagc	gaaaatacgt	ctcaaaaaaa	caaaaacaaa	caacaaaggc	acgtattaaa	128400
tacqaacata	aatatttaca	aattatactq	aataagttct	catgtttatt	atttgcttgt	128460
ccagttacaa	acttttcctt	cqtaqaatta	gaaatataaa	taataaacat	gagaactcat	128520
tcagtataat	taataattat	taaatqtaaa	taaaaacatc	tatgtacaat	taggcattta	128580
tttaagaatt	atttgaaaaa	aaaacaatgt	ggaaacagat	attttqatat	attectagte	128640
attgaaattg	ataatgttct	trtgaagagt	aaagtgacca	tatatattaa	agttaaaatt	128700
tangerange	atcacacgcc	taataaatta	tettaaggaa	atcagtttga	aagtaaaatc	128760
caactcagca	caaagacttt	aggregagera	2522223344	assatroagt	ttcaaattat	128820
aatatatgca	Caaagacttt	adcatttatt	ttaatataaa	attatata	tactttcctc	128880
atcctatgga	ctattttctg	Cladadayia	ccaacaccaa	atteraction	tacticities	128940
acaaatattt	tgggggagaa	aacccaacaa	aactacatge	actgcaattt		
tttttttta	gacagtettg	ctccagcgtc	caggetggag	tgcagtggtg	caatetegge	129000
tcactgcaac	ctccatctcc	caggttcaag	caattctcct	gcctcaggcc	tecegagtag	129060
ctgggattac	aggcgctcac	caccatgcct	agctaatttt	tatagttttt	agtagagatg	129120
gggtttcatc	atgttggcca	ggctggtctt	gaactcctgg	tctcaagtga	teegtetgee	129180
teggeeteet	agagtgctga	gattacaggt	qtaaqccact	gcacccagcc	ttatgcatta	129240
taattttaat	ttqtaaactq	tacaaaqqqa	taatacttgt	agtacaacaa	gaagtaaaaa	129300
catttottat	aggtagttaa	catttgtaac	cagtagaatt	ataggtaaaa	tttatttatt	129360
taaaacaqtt	ttagttggat	ttgatttcaa	ctttaaaata	atgcttttca	tctctatcag	129420
gtctttttgc	ctggcttttt	gtccagcaat	ctttattata	aatatttgaa	tgatctcatc	129480
cattcoottc	gaggagatga	atttctgggc	aggaacatat	cactaactac	tectagetet	129540
attaccata	ctgatgagtc	traccragge	agttctgaca	actetacate	tcaggtattg	129600
actgattggg	tctgccatta	accesses	catacacatc	ctttccttca	catcccagta	129660
actgattgtg	ttatttgtaa	3554544445	ataassass	aagataaaag	ccaggcacag	129720
acagatetta	statustasa	accccaagec	geggaaaaaa	tagacaaaag	acacqaqqtc	129780
rggeergrge	ctgtaatccc	agcactttgg	gaggeegegg	tatataataa	22252222	129840
aggaattcga	gaccagcctg	geegaeacgg	Lgaaacccca	cttccactaa	tanaganaa	129900
attagccggg	catggtggca	ggcacctgta	accetageta	cregggagge	Lgaggcagga	129960
gaatcgcttg	aacccaggag	gcagaggttg	caatgaacca	aaatcacgcc	actgeactee	
agcctgggtg	acaaagtgag	actgtgtctc	aaaaaaaaa	aaaaaagaga	gaaataaaat	130020
tagcctactt	actatcttct	aatcaaagca	tttgtggtaa	cttaaaatat	actgtattgt	130080
aaagtatcat	gctgtttcat	ttaggccatt	attctatttg	aatctgtggc	tgtttctctt	130140
aataaatcaa	gtaatatgga	atatattcat	agcctctgaa	gagctcttta	tgtaagtatt	130200
tatttaggat	actttttgta	aaataagtga	atgaattctt	aggtctcctt	tttttttctt	130260
ttcttgagac	agggtctcct	cqctqcaacc	tggaaattct	gggctcaaat	aatccaccca	130320
ccacagcete	ctgaatagct	gggactagag	gcatgcacca	ccacgcctgg	ctaatttgaa	130380
atttttttt	ggccaggcat	gatggttcac	gcctgtaatc	ccagcacttt	gggagaccga	130440
ggcaggcaga	tcacgaggtc	gggagatgga	gaccagcctg	gccaacgtgg	tgaaaccccg	130500
tctctactaa	aaatacaaaa	attagctggt	tatggtggct	catgcctgta	atcccagcta	130560
cttgggaggc	tgaggcagga	gaatggcttc	aaccaqqqaq	toggaggttg	cagtgagccg	130620
agatcacgcc	actgcactcc	tacataataa	cagagtgaga	ctccatctca	aaaaaaattt	130680
ttttttaaa	tgatggagtc	ttactatatt	actcagacta	gtcttgaacc	cctgacctca	130740
aataccacct	gcttcagcct	aagtttcttt	rttttttata	aagagacagg	atcitactat	130800
attaccaca	gtagtctcaa	actectaget	tcaagcagtc	ctcccacctt	ggcctctcaa	130860
griggicagg	ttacaggcgt	gaaccactac	ctataatott	gtgtttcact	caaggcttt	130920
agractagga	ttgcattacc	gaaccaccac	ctacaatgcc	ttgacctttt	ttaggttttt	130980
tgatttcgtt	tcatatgtta	gegeeacace	gregication	aaaatcacac	aaagtagaat	131040
tggagtgctt	cacacycca	aaccatactt	gattettet	adaattatat	taaactcaca	131100
atcctaagac	aagaaatcta	aggaggcata	aagaagttaa	ctggttttat	caaacccaca	131160
cagtaaatga	tagagccaga	aatattcccc	ttctagtgtt	cccaccatc	agettaatgt	1311220
agcataataa	ttttctaatt	actgttgaca	aacaaacaac	cctttgaatt	tttaatattg	
ggccttggat	aaattttcct	aatttgtaag	agagtattat	cgtattgcca	tttacaaage	131280
tctcctgagt	atctttttct	tctgttaagt	ttacctagga	gataaactgc	tgagtatggt	131340
tgccattttg	gttttttgat	ataggttaga	atgtcttggt	tttttttt	ttttttttg	131400
atttttatta	ttgtcattgt	ttgagacagc	atcttgctct	gtcgcccagg	ctggagtgca	131460
atggcacgat	cqtqqctcac	tqcaacctcc	acctcccggg	ttcaagcaat	teteetgeet	131520
cagetteetg	agtagctggg	attacaggca	tgtgcaacca	cacctggcta	atttttgtgt	131580
ttttagtaga	gaaggggttt	caccatgttg	gtcaggctgg	tattgaactg	ctgacctcat	131640
gatccacctq	cctcggcctc	ccaaaqtqct	gggattgcag	gcatgagcca	ctgcacctgg	131700
ctgaatgtct	tgtttttgat	taggcactta	agaaaggcct	aggtactaac	cataaaatat	131760
atttttatac	cttttgttga	tactatatat	atagaaaact	gcacttatca	taaccttaga	131820
	aatgttcaca					131880
cacceagaag					3 · · · · · · · · · · · · · · · · · · ·	_

WO 01/27857 PCT/US00/28413

cagcattatc	agcccctcta	gaagccctct	tgggcccctt	ccattcactg	tecttettgt	131940
caccagggta	gctactatcc	tgacttttga	tggcatagat	tagcattacc	tgttcttgtc	132000
attttataaa	taaaaccata	ctatatattc	ttttcttgta	cagctttatt	gtgctaattc	132060
acatttacat	catacaattc	agtggttttt	atatggtcac	agagttaggt	aaccattacc	132120
	tagaacattt	ttttcactcc	agatagaaac	cccctttact	taaactccaa	132180
atccccact	ccaccagccc	taggcagcca	ctagtctact	ttttatctct	atagagacaa	132240
tagatttgct	tattctggac	atttcataaa	catogaacco	tatattatqt	ggtcttttgt	132300
taccaactat	ctttcactta	gcatcatgtg	ttcaaaagag	catcatotta		132360
gcatgtatca	gaattttatt	cctcattatg	occaaatatc	ccattgcaag	gatttatgac	132420
attttattta	aattgtaccc	teetttetee	catttatcaa	taatoctact	gtgaccattt	132480
atatacaaat	ttttgtgtgg	atacaggttt	tctttttatt	tttaaatttq	aggtggagtc	132540
ttactctata	gcccaggctg	gagtgcagtg	gcacaatctc	gactcactac	aacctctgtc	132600
tectagatte	aagcagttct	cctacctcaa	cctcccgagt	atctgggact	ataggcacgc	132660
2002999000	ccagctaatt	ttttagtaga	gatogggttt	caccatgttg	accaatctaa	132720
tetecace	ttgacctcaa	atastcacc	catctcggcc	teccaaagtg	ctgggattac	132780
	cactatgccc					132840
tagaggrage	gagtcaagag	ggccgcggcc	aacttattga	aaaactgcca	gattgttttc	132900
cagaaccgcc	gcaccatttt	gcaaccccca	cagcagtgta	tgagttttac	agcttctcca	132960
cyaaaayycc	gaacttatta	tetatttaac	tatttttaaa	aatgatagtc	attccaataa	133020
cattleatty	agtgtggttt	ttagaettet	ctaataaata	atgatgttga	gcatcttttc	133080
grectacted	tggcctttgt	totacette	gaaaatgtt	tattcaaatc	ctttggccat	133140
tttttt	attttattt	atttatttt	ttttgagacc	aagtctcact	ctatcaacca	133200
tactet	caatggtgtg	atctatttt	actocaacct	ccacctccta	tattcaaata	133260
ggctggagta	ctcagcctcc	casatageta	ggattagatt	traggracet	accaacatac	133320
accentent	tttgtatttt	tactactcac	aggatttcac	catgttagcc	aggetogtea	133380
cgggctgatt	acctcaggtg	atetagetae	ctaggettee	caagetagee	aggetggtoa	133440
caaactcctg	tgggcccagc	attegette	tttttttt	ttttttta	gaaccacags	133500
egegagecat	gcccaggctg	gagagagaga	ccacaatctt	aactcactac	aacctctcc	133560
testesst	aagcgatttt	cetacetesa	cctccccact	ageteggett	acadatacct	133620
coccaggere	ccagctaact	tttgtttt	ttttagagag	ageegggaee	catottoocc	133680
accaccacac	caactcctga	cctgtattt	tecaectece	ttaacctccc	gaagtgctgg	133740
aggetggtet	caactectga	ccccaggcga	aattttataa	ttatattccc	carretrate	133800
gattacegge	atgagetace	aggeeeagee	aatttcccca	ccaeactgtc	caggetggte	133860
ccaaaccccc	gggttcaagt	gattettet	cottettaaa	ttagattgcc	tttttatatt	133920
gegegageea	ccttgctcag agtcctttat	atattatage	tanatatasa	ttatcaaatt	atattatttc	133980
gagtttcagg	agteettat	tanattatat	ttaatgtccc	ttttaaaaa	acaccacttt	134040
caggtattt	cttcattctg	tgagttgttt	aggretate	tetectecce	aggegggeete	134100
tgtttgtttg	tttgtttgtt	coctcaagac	tanacttact	cccgccgccc	gatectetta	134160
geaguggeae	aatcacagct	cactgccacc	atagagagta	teacacccac	cetette	134220
etteageete	ctgaatagct	agggecatag	acacacacca	teteaaacte	taggeteaaa	134280
ctgtttgtag	agacagatct	tactgtgttg	ctccaagttgg	aggtataaga	caggeteaaa	134340
gtgattetee	cacctctgcc	ccccagageg	cegggaccac	aggractage	tacaegeaae	134400
etgtetttte	actattaata	gracettett	ttacatttt	actacacaca	grattcacc	134460
ggcacccacc	accatgcctg	gctaatttt	ceteseetes	tteacctccc	atggcctccc	134520
atgiteacce	ggctggtctt	gaacteetga	cettaggtga	anastattee	cttcttaaca	134580
aaagtgctgg	gattacaggc	gtgagetaet	geacceggee	tattatata	ttettteatt	134640
grattgrett	ctaatttgtg	aacatggatg	catetteatg	ctectteest	catttattcc	134700
tcagcagaat	tttgtagttt	ccagagtaga	agectttcac	ettettigggt	tttcatttc	134760
tatgttttaa	gttcttttcg	attecattat	aaatagaatt	tagatattaa	tettecaact	134700
agattgtttg	atgagagagc	atagaaatac	aagtgattt	ttatatatta	tttaggattt	134880
teaactttga	taaatctgat	Egitagetet	aatagtttt	tttttt	ctaggatte	134940
tcaatatata	agatcatgtc	atttatggat	agagatagtt	tetetetete	ccagaaccca	135000
cagagcaatg	atgagtagaa	grggcagaag	tastatta	tatacattt	castasatoo	135060
agggaaagct	ttcagtttca	teacttaata	igatyttägg	tttaaggtttt	ttttttt	135120
cctttttcag	attcaggaat	ttccctatca	cccccgattt	tatatasata	tabateatea	135120
ttaaatcatg	aaagggtgtt	gaatattgtc	acgeeette	Lgcaccagca	gatgttaace	135160
tatggatttt	gggttttatt	cegeegatgt	gaaatattaa	cigalitica	ttcaatatco	135340
caaccttgca	tacctgagat	gaatctcact	tggtcatggt	grataarett	ttaagataag	135360
tgctggattc	catttactgg	tatttgttg	aagattttgt	accigaacgc	taayataaC	135360
atttacactc	tatcagaaat	gaattgacca	Laaatgtgag	agegraticg	tategasata	135420
attetettee	attccaaaga	tagacataca	tccgtctgta	rataratat	Lacyccayta	T33490

```
ccatactete tigattaeta tigettigta ataagittig aaateagaaa giataaatga
                                                                                135540
gattttggta totgagtaac agtoctoata gaattagttg ggaaatatto cotottaatt
                                                                                135600
ctggtccctc tttcttttt gtttaactgt gtatcttgga gattgttcct tctcaacaca
                                                                                135660
tgagageege tttccctace etcccaccee tgctatagag aggtetataa gtgtetgtte
                                                                                135720
aattatttta tttacttaac ctattactta gtcggggaca ttaagcttgt ttatgtcttt
                                                                                135780
tattttaaac aatgctgcag tgaataatct tgtatataag tcattttcca tcaatataag tctctctgta actgaatttt tagaagtgga atttctaggt caacctatgg ctctgtattt
                                                                                135840
                                                                                135900
cacaaaaata ccaattctgg tttttcttgt ggaggtgggg agtaggaggt agaatgctgg aggagaactt gctgtactca gctggctagt cattttagaa aggtttcctt agcttcttt
                                                                                135960
                                                                                136020
tgtcatatgg cctcaccaag aatcaaaaac attcctattt accctgtaaa catggggctt
                                                                                136080
tactacccaa gatacatatt tetggatgta tgacagettt teatattgaa gaaataatge
                                                                                136140
tgtgagtaca gcacatttgt tggaacttag gtcgttaaga atgtcttata aattcatacattatacattt tattttatt tatttttag tttttgatac agagtcttcc tctgtcgccc
                                                                                136200
                                                                                136260
aggecagegt geagtggtac aatettgget cactgegace tecateteet gggetcaagt
                                                                                136320
gatteteatg teteageete eagagtaget atggttaeag geatgeacea ceatgeeegg etaatttttt tattttagt agaaactggg ttteaceata ttgaceatge tggeetegaa
                                                                                136380
                                                                                136440
                                                                                136500
ctottggcot caagtgateg geotgeotea geoteceaaa gtgetgggat cettgtattg
ggtaaaagat gaatattgag ggctgcatgg tggctcatac ctgtaatccc agcactttct
gagactgagg tgggaggagt cctggagccc aggagggtga ggctgcagtg agttgtgatc
                                                                                136560
                                                                                136620
gegecattge actteaacet aggaattata ggetteagte actgtgeeeg geatgtacat
                                                                                136680
                                                                                136740
tttaatattg tgctttcctc ttttagctat agtatgaggt tacatttcag agtcattgtt
gttaagcatc ttaatagtga tgaggttgag tgaaagttac ttctatttca aacactgaag
                                                                                136800
aaaattttgt adaaatctgt cacattccaa gcccaggact gattgtttca tatacttcta
                                                                                136860
attttacaat ttctattgta gtccagtgtg aaaaaagcca gtattaaaat actgaaaaat
                                                                                136920
tttgatgaag cgataattgt ggatgcggca agtctggatc cagaatcttt atatcaacgg
                                                                                136980
acatatgccg ggtaagctta gctcatgcct agaattttta caagtgtaaa taactttgca
                                                                                137040
tottttaaat titttaatta aattttacat tittttotaa totattatta tatgoocaga
                                                                                137100
actiticacti agagtgigca gialiaatgig giggitaagi alaaaggcic iggagtgact teetgggitt taatetigge teigecatti aliggeagee gelaacetei iggialeta
                                                                                137160
                                                                                137220
                                                                                137280
gtttcttcat ctgtaaaatg agaataataa agtgaaaaga tgccaacatc atttactctg
                                                                                137340
ggctgcataa ctgatacttg gaaaaagtat tcctttgagt ttaagaatta agttggttat
                                                                                137400
tcattttage ttgtaataaa aagatagtga ttcataggat atgccaetta etgaaattta
ccacagatec aateataaaa teaetttete tteeetaaag atagettgat taacatgtaa aggtgtgtaa aggettgatt acactaceet gateegtace ccagtteeca geageaceat
                                                                                137460
                                                                                137520
                                                                                137580
gaaaaaggga titcaacata tttaattact ttcagtagaa agtaacagtg gtaggccagg
                                                                                137640
cqcagtggct cacacctgta atcccagcac tttgggaggc cgaggtgggc ggatcacgag
gtcaggagat tgagaccatc ctggctaaca cgatgaaacc ccgtctctac taaaaataca
                                                                                137700
aaaaattage egggeatggt ggeaggeace tgtagteeea getaettggg aggetgagae
                                                                                137760
                                                                                137820
aggagaatgg cgtgagcccg ggaggcggag cttgcagtga gcttagattg tgccactgca
                                                                                137880
ctccagcctg cgcagtggag cgagactctt gtctcaaaaa aaaagaaagt aacagtggta
ttgggagact gaggagccta gaaagtactt gaaggaagta aaaggtttgt ttgaccacat
                                                                                137940
                                                                                138000
tqtatttgga aagccagctt tttcagctgt gtcagctttg tgtagtgatt tttagttctt
cttttagaaa ataacggaca aggccgggca cggtggctca cgcctgtaat cccaccactt
                                                                                138060
tgggaggeeg agaegggegg attacetgat etcaggagtt egagaceage etgggeaaca
                                                                                138120
tggtgaaacc ccgtctctac taaaatacaa aaagttagcc gggcgtggtg gcgtgtgcct
                                                                                138180
gtagtcccag ctactccgga ggctgaggca ggagaattgc ttgaacccgg gaggcggagg
                                                                                138240
ttgcagtgag ccaagatcac accattgcac tgcagcctgc gcgacagagt aagactctgt
                                                                                138300
ctcaaaaaat aataataaaa taaaaaagaa tggacagtaa acctaaatga gttcattccc
                                                                                138360
aaagatgatg ttattettaa gggatggtte atttatttaa gacettacat aaagtetate
                                                                                138420
                                                                                138480
aattgogtga tttttcactt ctgtaattgt gtgtatgtat aatgtaaata tatatgtttt
tgttttgttt tggttttttg agacggagtc tcgctctgtt gctcaggctg gaatgcagtg gtgcaatotc agctctctgc aacctctgtc tcccaggttc aagcgtttct tctgcctcat
                                                                                138540
                                                                                138600
cctccaagt agctgggact acaggcacgt gccaccacgc ccggctaatt ttttgtattt ttagtagaga tggggtttca ccgtgttagc caggatggtc tcaatctcct gacctcgtga
                                                                                138660
                                                                                138720
                                                                                138780
tecaceage tiggettece anagigtige tattacagge atgagecace acaceagea
tgtatttttt aaatgtataa aatgaagcag aaaagagaaa tgataatttt tcttcatctt
                                                                                138840
gaaagattat cttcaccagg cgcagtggct cacacttgta atcccagcac tttgggaggc
                                                                                138900
cteggeagge ggeteactig agttegaaac eageetggee gacatggtga aacteegtet ctactaaaaa taaataaata aagatggttt taatatatgt tttagtttta tgattttage
                                                                                138960
                                                                                139020
                                                                                139080
atctttctga aatttttctc aaggcaagta aatttgtatc agttggtata ttggtaccca
```

WO 01/27857 PCT/US00/28413

```
tctatgaaat aacttattag gaagatatct ctaaaataag atcactttgc ctaaaataaa
                                                                                139140
ctgatatatt gatgttcaca gaatttttct tttaaccgac ttgataaatg cattattctt
                                                                                139200
gacgtcaagt gatccacctt cctcagcctc ccaaagtgct gggattacac acatgagcca
                                                                                139260
ccgcacctgg cattattett ataaaaggtt aaatttetag ttaagtttaa tgteetettt
                                                                                139320
gttcatgtac cattgcttat tttcttccct tectactcac agtaatcatt cttatggtat
                                                                                139380
geactitigt tigettatit tratgraatt gatattacge tecattetgt acgitgtact
                                                                                139440
ttcattcaca gtgagttttg gacattccta tgttcatcta tacagactta cttcatttta
                                                                                139500
actacactgt agtattccgt atgtaatatt tactataact catcactgta gcagagcatc
                                                                                139560
tcatagtgta tgtattactg ttttgccatt ttggtatcaa tgagtattta agtcatttgc
                                                                                139620
                                                                                139680
agtttttccc tcttataccc agtattacag aggatctctt tttatatgct tctttgtacc
                                                                                139740
aagaggcaga ttaaaaaatt titttttgaa aaaatttttg aaaaaaaatg aaatgaagtc
tcactatgtt gcccaggctg gtctcaaact cctaggctca agcaatcctt ccatcttggc
                                                                                139800
ctcccaaagt gctggggtta caggcatgag ccaccatgcc tggcctacat tttaaatttt gatagctctt acaatttact ttgtaaagta tctgcatcat tttatgttct caccagtctt
                                                                                139860
                                                                                139920
                                                                                139980
taataagaat acttcatact tttggctgga cacagtggct cacgcctgta atcccagcac
tttqqqaggc cgaggcgggc agatcaagag atcgagacca ccctggccaa tatggtgaaa
                                                                                140040
ccctgtctct actaaaaata caaaaattag ctgggcgtgg tggcgcaccc gtagtcccag
                                                                                140100
ctactogaga ggotgagaca ggagaatcac ttgaaccogg gaggtggagg ttgcagtgaa cttagatcac accactgcac tccagcotag caacagagtg agactotgto tcaaaaaaaa
                                                                                140160
                                                                                140220
                                                                                140280
aaaagaatac ttcagactta atttttttc cagtcttaag tgtttgctaa tgagattgag
tttcttttgg tatgtctctt gattgttcag gttttttctt ttatgaattg actgttcatc tcttttcac attattctg ttgggtgatt ttattagtga cttgttaaaa ttctgtatat
                                                                                140340
                                                                                140400
tttttcagca tgacacttca ttattcaaaa aaaaaaaaag attctctatg tttctcgata
                                                                                140460
ctaatcattg gttggtaata ccttaaaaat aagaccctta ctgtattttt tgctttttt
                                                                                140520
                                                                                140580
ttttttttt ttttttt tttgagatag agtettgete tgttgeccag getggagtge
aatggtatga totoggotot cagotoactg caactgoaac ototacotoc otgtttcaag
                                                                                140640
caatteteet geettageet cecaagtage tgggattaca ggeatecace accacaceca
                                                                                140700
gctaattttt gtatttttag tagagacagg gtttcaccat gttggccagg ctggtctcaa actactggcc tcaagtgatc cgcctgcctc ggcatcccaa agtactggga ttacaggcat
                                                                                140760
                                                                                140820
gagccacagt gcctagccac titttgcttt ttaactttgt titatagtac tatagttta
                                                                                140880
gtataaacag atgtatgtat acacacaact atggctttat aatatgtttc agtcattgtt
                                                                                140940
agagcaaggc ctaccttttg ggtgcttctt ttacaaaatt gtcttggcta ttcttgtgct
tttttctta tttgtgaatt ttagaattgt gaattacctg ttgactcacc atgtttgta
aactgaggat tttgaatgga attgcactca attaaagatt atcttgcttt ctgtgcagca
                                                                                141000
                                                                                141060
                                                                                141120
                                                                                141180
atgttttatt tcaaataatc cctactttaa attacttagg atagctataa attgtgtttc
tggctttcta gatttagatg aaacgcttta aattgattgt tttctcctaa atttaaaact gattgttaga agttaaagtc ttctgttcat tcttatttag gaagatgaca tttggaagag
                                                                                141240
                                                                                141300
tcagtgactt ggggcaattc atccgagaat ctgagcctga acctgatgta aggaaatcaa
                                                                                141420
aaggittgtg gigittitat acticatatt aagcetttae teacattagt gattgaeigt
                                                                                141480
aagtcaaaga ccacttaagg tttaaactgt ttattttgta aagtaaccac tgtatctttc
accttgtgtt tatagtcaga agtaagtaca agggcttcct gtagtcacat ctttatgcaa
                                                                                141540
totoctotga atcaaaagtt agtgaacttg otttgccact ccagaaggca catgaatatg
                                                                                141600
aaaaagcatt gtctattttc ttatttaatg gcaaaatacc cgacctaagt tggacttaat gtttgagacc gtttatttta ttaaattata ttttttctct tttcttttt ttttttgaga
                                                                                141720
                                                                                141780
cagttettge tetgteacce agaceggagt geagtggtet gacegeacet caetgeaace
totgetteet aggiteaage gattiteetg ceteateete etgagtaget gggaetacaa
                                                                                141840
gtgcgcacca ccacacctgg ctaatttttg tatttttagc agagatgagg tttcaccacg
                                                                                141900
ttggctaggc tggtctcata ctcctgacct caagcaatcc atccgccttg gcttcccaaa gtgctgggat tacaagtgtg agccaccatg cctggcctta ttaaattatt tttattaaat
                                                                                141960
                                                                                142020
                                                                                142080
ttootoaaga ttgatgaaag taatgaaata taaaagtaat gaaatatatg tggaaaatag
actggattaa gaaaatgtgg cacatataca ccatggatac tatgcagcca taaaaaagga
                                                                                142140
tgagttcatg tcctttgtag ggacatggat gaagctggaa accatcattc tgagcaaact
                                                                                142200
gtotcaagga tagaaaacca aacaccgcat gototcacto ataggtggga attgaacaat
                                                                                142260
                                                                                142320
gagaacactt ggacacaggg tggggaacat cacacgctgg ggcctgtcgt ggggtggggg
gctgggggag gaatagcatt aggagatata cctaatataa atgacgagtt aatgggtgca
                                                                                142380
gcacaccaac atggtacatg tatacatatg taacaaagct gcacgttgtg cacatgtacc ctagaactta aagtataata aatttaaaaa aaataaatat atgtggaaaa tattaatagg
                                                                                142440
                                                                                142500
                                                                                142560
tcaaaattca aattgttcat`ttaatcagaa gagtagttta gtcaaatcca agggttagac
aacagaaatc ttttttgtca agtgcattct ttgtgactga tttcattttc ttcctggttt
                                                                                142620
                                                                                142680
acacaggaag atttcagaaa caaatgtgga tocgtgacag atggtatcta gaagttttta
```

```
gtttggttga attgacagta ttttattgag taaaagatac taatttttgt aagaagaaaa
                                                                                 142740
                                                                                 142800
attcaatttt gataagtatg tttaagatta agagctattg gccaggcgct gtggctcatg
cctgtaatcc tagcactttg ggaagctgga gcaggtgggt cacgaggtca agagattgag accatcctgg ccaacatggt gaaaccctgt ctctactaaa ttagccaggc gtggtggcac
                                                                                 142860
                                                                                 142920
atgeotytge accegeotee gygtttaage gateetactg ceteaggete etgagtaget
                                                                                 142980
gggattacag gcgccatggc taatttttgc atttttagta gagacagggt ttcactacat
                                                                                 143040
                                                                                 143100
 tggccagget ggtetggtet caaacteetg accteaggtg atetgeeege ettageetee
caaagtgctg ggattacagg catgattcac catgtctggc catttatctt attttcttt
ttttttttt ttttgtttga gacggagtct tgctgtgtcg cccagagctg gagtgcaatg
                                                                                 143160
                                                                                 143220
gtgcgatctc agctcactgc aacctctgcc teetgggttc aagcaattct cetgcetcag tettccaagt agctgggatt acaggegegt gecaccacat ctagetaatt tttgtatttt
                                                                                 143280
                                                                                 143340
tagtagagac agggtttcac catgttggcc aggctggtct cggaactcct gacctcgtaa
                                                                                 143400
tetgeceace teggeeteec aaagtgetga gattacaagt gtgagecact gtgeceagec atettattt etttetttt ttttgteggg tgggagggg acagagteta getetgtege
                                                                                 143460
                                                                                 143520
caggettgge teactgeaac etetgeeece caggttetag caattattet geeteageet
                                                                                 143580
cccaagtagc tgggattata ggcacctgcc accacgcctg gctaattttt tgttattttt
                                                                                 143640
agtagagatg gggttttgct atgttgacca tgctggcctc aagtgatccg cccaccttgg
                                                                                 143700
cctcccaaag tactgggctt acaggcgtga gcttgtattg ggtaaaagaa caatattggg ggctgcatgg tggttcatac ctgtaatctg agcactttgt gagactgaga tggaaggagt
                                                                                 143760
                                                                                 143820
gttggageee aggagggtga ggetgegget geagtgaatt gtgateaege cattgeaett
ceaectaggt aatggageaa gaceatgtet etaaaaaaca aaacacaatt tttttaagga
                                                                                 143880
                                                                                 143940
atactgggaa gaggtcagtg gtggttttag aacagaggaa gtgccagatg acctttgtga ggcattggcc aggaagaact ctacagtgtc tttaggtagc ttctgtccat aaggataatg gggtctcctc cccagtatta atagaaaatc tctgagctgt ttttttttgt ttgtttgttt
                                                                                 144000
                                                                                 144060
                                                                                 144120
                                                                                 144180
 tgtttttttt tootgagatg gagtototot otgtoggoca ggotggagtg otgtggogog
atettggete actgeaaget etgeeteeca ggiteacace atteteetge eteageetee
                                                                                 144240
caagtagetg ggactacagg tgtccaccac cacgeccage taattttttg ttattttag
                                                                                 144300
tagagatggg gtttcaccat gtcagccagg atggtctcga tctcctgacc tcgtgatccg ctcgcctctg ccttgcaaag tgctggagtt acaggcgtga gccaccgtgc ctggcctggt
                                                                                 144360
                                                                                 144420
144480
gtegeeeggg etggagtgta gtggeaegat gteggeteae tgeaagetet geetgeeagg
                                                                                 144540
ttcaagccat totootgoot cagootootg agtagcaggg accacaggeg otogcoacca
                                                                                 144600
cgcccggcta attittigta tittitagaag agacggggtt tcaccgcatt agccaggatg gtctcgatct cctgatgtcg tgatccgcc acctcggcct cccaaagtgc tgggattaca
                                                                                 144660
                                                                                 144720
144780
gacageteat gaagaagtge teetgettea tatgtatatg tgttageata gtgttaacat
                                                                                 144840
agcalaggtg ltcggtgtlt gcaglttctg tttgttttal algaaltaag gtgtattatg
                                                                                 144900
agcagttgaa gatatatagg aaattttttc ccaaaccact atctctgctc gttctattca
ttcagtctgt ttatgttatt ccttcattca ttcattttat agaacagtgg agtgcctact
                                                                                 144960
                                                                                 145020
gtatgcatct attgttctgg gtcctgggga agaaaacaaa gttcctgctt tcatggaact
                                                                                 145080
tacattatat tggcggagac agtaacagac aaacaaatgt agcctgtgta catgtgttac
                                                                                 145140
atgaaaagca gggtaggggg ctgggagaga gtagtaggga gtgctatttt cgaggtggtt
                                                                                 145200
gtcaggaaag gcctcactga ggaggtggca ttttgagtag acctgagcgc agcgggggcg
                                                                                 145260
                                                                                 145320
taagcccagg cagcatgtgg aggaagagtg ttcttggtga aaggaacaag gatagaggcc
                                                                                 145380
cgaagctaga gagctcagca tgatcaagga acagcaagcc ccgtgtggct ggaatggagt
gagcaaagga atgagcagta gaaggtgagt gagttgggag gtcaccagag accatggcaa
                                                                                 145440
                                                                                 145500
ggacttgaaa gtgtcaggga cacattggaa gttggagcag ggaaatgatg ggatttatgt
tttgtttttg ttttatgttt agtgttttta agggattgct ctatcagcta tttggaaaat
                                                                                 145560
ttagtgtagg gcttcaagaa gagaagcaga gaaacaacat tcttgccata gtcatagtct
                                                                                 145620
                                                                                 145680
aagtaaggga tgatggtggt gtggattagg ctggtagtgg aagaccagtc cagttcgggt
tgtatttgaa ggtagaggca aaaagattat atttctacca gcaagcccat ctatgaagtt
                                                                                 145740
acttgtatta ttaatttaat tgagacatgo coacataaac taataaatag gaatttotgo
                                                                                 145800
                                                                                 145860
agttiggtta aacaccctg tatatcctig ticttettt agttgtecag atgtetettt
                                                                                 145920
aagtcaagta ttttttggtg gtgtaggagc ctagagattg aatttattca cccaaaaggc
atttgagtga ttactatgtg ccaggcacta tgctgaatgc caaggatgta aataagaggg
                                                                                 145980
cgtagtetea gtetgttta etceagettg gtteettttt aatgaceetg aettgttaag eatateagtt atcetacaga atgtttaate ttetgtaett teetggttgt gttatttage
                                                                                 146040
                                                                                 146100
ttatttctct ttccttgaca tttcttgtaa actggaagtt acacctatag tcttgatgat
                                                                                 146160
tegtgttaca cattttägat tagaacacat catgtgttgt atatggtgtt tttgaaagee
                                                                                 146220
tctctgtata ttggtctgta cattaaaatg ttgcctgaat ggatacacat aaaatttaac
                                                                                 146280
```

```
agtgattaca ttagagatga gaagaaagag gtgcctttta cttttcaata taccttttcc
                                                                            146400
totgettttt gaactitett geectatgea taegttattg ettaateate caecteatet
cttcccctgt ggctttctgt tgcatttgga atgaaatcta gcctctttgc tgttacctgt
                                                                            146460
ggatgteect tgetggeete tateacetta etttgaacea eteettteat ggaetgaget
                                                                            146520
ctcattggac tatcitttat tettitgetg aagtitette aetttgagtg cetetgeagt
                                                                            146580
                                                                            146640
tgctatttca tggctgtggc aagccctgcc atggctttca tgcaaggatg gttcctcctt
ctcatctcaa tattatctct tcagagaggg accttcccaa ctccgatgat ctaaaatcct ttgtatatac cactcactac cacttcttt tttttttt
                                                                            146700
                                                                            146760
ttttttttt gagatagggt cttgctctgt tgcccaggct ggaatcacga ctcactgcag
                                                                            146820
                                                                            146880
cctcatcttc ttgggctcaa atgatcctct cacctcagcc tctcgagtag ctggaactgc
                                                                            146940
aggeacacae caccataett ggettattat tttaettttt gtagagaeag ggttteacea
aggotggtot caagotootg cogcaagoaa tocacatoto toagootooc aaagtattgg
                                                                            147000
gattatagga gtgagccact actcctggcc tattttctta ttcactgtct aaaattatct tgttcattta tttacatact tgtttatagc ttatttctca gctggacatg gtgcctcaca
                                                                            147060
                                                                            147120
                                                                            147180
cctgtaatct caatactttg ggaggctggg ttggagaatt ggttgagccc aggacttcaa
gaccagectg ggcaacaaag tgagaccetg tetataaaaa attgtttaaa aattagetgg geatggtgge acatgcetgt ggteceaget acttgggagg cagaggtggg agaategett
                                                                            147240
                                                                            147300
gggcccagga ggttgaggcg acggtgagcc atgattgtgc cactgcactc tagcctagtg
                                                                            147360
                                                                            147420
acagagtgag accatgtgtc taaaaagtaa ataaaaatag tttctctttc atgactagaa
tattacetet atgtgggeag ggagtttgte tatactattt ggeactatat tteetgatte tgaaattatg cetageacat ggtaagtaet cettaaatat ttattgactg aattattaa
                                                                            147480
                                                                            147540
tacttaagaa tttcatttgg gattatctga gtggtaagat tacggattat atttatgtaa gaaaaaatca ttttttaaac ttggttgccc tttgccacac tgacatagac actaagtttt
                                                                            147600
                                                                            147660
cttagecaga ttaetteega ggataeteae agaggeeatt etetteteaa teeceaaata
                                                                            147720
attgatattt cttagcactt tcaagctaat gcaattetta gatgatgtat etgtgtatat
                                                                            147780
catatectea tectacaaat gtagaaattg aagtetggge acagtggete teacetgtaa tetcageagt tegggaggee aaggegageg gateactgag gacaagagtt aagaceagee
                                                                            147840
                                                                            147900
tggccaacat ggtaaageet tgeetetatt aaaaatacaa caattaggge egggegtggt
                                                                            147960
ggctcacgcc tataatccca gcacgttggg aggccaaggc aggcagatca cgaggtcagg
                                                                            148020
                                                                            148080
agttogagac catootggot aacacagtga aaccocatot otactaaaaa tacaaaaaat
tagocaggoa tggtggcacg ogottgtagt occagotato gggaggotga ggcaggtgaa
                                                                            148140
tecettgaac cegggaggeg gaggttgeaa tgagetgaga ttgeaceget gaacteeage
                                                                            148200
148260
ggtggcgggt acgagtacct gtaatcccag ctactaggga ggctgaggga ggagaatcac ttaaacccag gaggtggagt ttgcagcggg ctgataatgc accactacat tccagcctgg
                                                                            148320
                                                                            148380
gcaacagagt gagactctgt cttaaaaaaa aaaaaaagaa agaaagaaat tgaggaatgt
                                                                            148440
ggagattgtg gtctgtgatt tgttaggaat cacacagcag gttagtagca actacagggc
                                                                            148500
tttggttcag aataccacet tgacaatggt ttgtttacag ttcggctccc cttcctctgc ctttctctcc ttccttattg agggcagctg gaaagaattt tcatcattta ctagcctata
                                                                            148560
                                                                            148680
getttaattt gagttttgaa accttgataa tagageacag aggaaaagae tgagtttet
ttttttgaga cagtettget etatggeeca ggetggagtg cagtgacaec ateteagetg
                                                                            148740
gttgcaacct ctgcctccca ggttcaagca attctgcctc agcctctcga gtagctgaga
                                                                            148800
ttacaggcac gtgtcaccac gcccagctaa ttttctgttt ttgtttcgtt ttgtttttt
                                                                            148860
ctgagatgga gtcttgctct gtcacccagg ctggagtgca gtggtgcgat gttggctcac
                                                                            148920
                                                                            148980
tcaaacctct gtctcctggg ttcaagcaat tcttctgcct cagcctcccc agtagctggg
actacaggta cgtgccacca tccctagttc atttttgtat gtttagtaga gatggggttt
                                                                            149040
cactatgttg accaggetgg tetegaacte etgateteag gtgatetaet egteteagtt
                                                                            149100
toccaaagtg otgggattat tggcacacgo otatttttgt atttttagta gagacggggt
                                                                            149160
                                                                            149220
ttcaccatgt tggttagact ggtctcaaac ttctgacctc aagtgatttg cccgcccag
                                                                            149280
cctcccaaag tgctgggatt acaggcgtga gccaccgtgc ccagccaaga ttgagttttg
aaaagagcct tctgagatta tgagaagggc aagcaagata acttaagaag ttacattaaa
                                                                            149340
atcatctaag agacagtgta acaagaagga attgtaaaat gatgttatga gcacgtgccc
                                                                            149400
aatgtagtgg caatcccttg tgcttcgata cattggtggg agacaaaact gtacttaaat
                                                                            149460
                                                                            149520
tgataaatcc cttacatgtc attttaagga gcttagactg actcccatca tgtagacatc
149580
ttgttgccga ggctggagtg caatggcgtg atctcggctc accacaacct ccacctccca
                                                                            149640
                                                                            149700
ggttcaagca attotootgo otcagootoo ogagtagotg ggattacago catgoaccac
                                                                            149760
cacgcctggc taattttgta tttttagtag agacggggtt tctccatgtt gtggctggtc
togaactoot gacctoaggt gatcotoocg cotoagcoac coaaagttot gaaattacag gogtgagooa cogogocoag cocagagatt totaaacaga gttotaacca gatgotttto
                                                                            149820
                                                                            149880
```

WO 01/27857

cctgtcagta	gaatgagaat	gaattggagg	tgggagagac	tggcatgagg	gacaccagtc	149940
aqccaqtqqa	attagctggt	aatgttgata	ggagaagaaa	aagattcaaa	gttaggtagt	150000
ggtagcaaga	attagaggga	aggtcggatt	tatgatatgt	ccaaggttga	attctaaggt	150060
gaaatttggt	ggcagatttc	atqtqtaaat	tgggaaggta	gattgagttt	ttttaacatg	150120
ggttttctaa	catqtcaata	qaqtqactct	gcaggggggc	ctgacgagag	aacagtgcat	150180
ggggtgattc	aacagccagt	tgagccttca	tgcagagcat	ttaacactgt	gactctgtag	150240
actictogtto	gcagtaaaat	ttcattaaac	caatatttaa	accettaggt	aataataaaa	150300
attgagggaa	aaggatccag	gttttgtatt	ttttatgaat	tcaqttattq	aattaaacaq	150360
gaccttgcct	caagaaataa	tctaccaaca	attaacttgt	tttaaaqcaa	agttaggaag	150420
tgagcatgtt	caaattatta	aataaaaaaa	taagetgtgt	atttcattca	tagaaataga	150480
ggctggccta	cttcggatga	trctcagcat	gtgattacag	atgtgggctt	atacatccta	150540
ggccggccca	gcgtactctg	acttagatag	agragagete	tttgaaactc	ttctctcacc	150600
cacctacttt	atatagacta	gereggarag	atgtagcagc	atactctotc	ttagaagccc	150660
ttttatatac	gagctggtct	ggaagtttg	aaaacataac	aaatgtgttg	gtgtctccca	150720
atotattoot	agattettae	ccaagaccat	tatecteett	aggatttagt	ttaattttat	150780
tttatttt	aatgtttgcc	acaagagcac	actacatott	agttettea	tcaagtgagg	150840
	aaaagtccag					150900
getegeaage	tadatyctct	gitatgtaag	caacacaacc	tateteease	aatgtaacat.	151020
teettagtgt	tttggggtat	Cacacaaaaa	agaatattta	tacciggaag	tttaaattaa	151020
taaataagag	cattgtggtg	gragragraa	cagugguucu	totoccttc	ettenneste	151140
agtetegete	tgttgcccag	gttggagtge	ageggeaega	cottageteg	ottcaacccc	
tgctcccagg	ttcaagcaat	tettetgeet	cagceteetg	agtagctggg	attataggca	151200
	tgcctggctg					151260
	cttgaactct					151320
	ggcatgaacc					151380
atgcatgaaa	tgtacattca	attttgtctt	tgtttactag	gatccatgtt	ctcacaagct	151440
atgaagaaat	gggtgcaagg	aaatactgat	gaggtaaatc	ctacctttag	gataaaaaga	151500
tttctgttta	taagtgccac	cctcatgtaa	gtgaggttta	aaattttcct	tttctttagg	151560
tcccatgttt	aagcagcatg	gcacatttat	gttctcttac	ccagaatgta	ccaagaaagg	151620
gtggtccctt	cttaacatct	aacaattgcc	tggtagtagc	agtgaaggta	tcttcagtca	151680
gaggctagga	ccactgaagg	atatacatgc	attcaagttt	ccatcagcca	gcaggcatca	151740
gtaatcagtg	tgtagatcaa	aagctcaaat	gtttccttcc	ccactggcag	ttttacttca	151800
agtagtggag	gcttgctttt	ttaatagtta	attaagtaca	ttgagagatg	ggaggtgaaa	151860
aaaggaaaat	gttttatttt	gaccatctaa	tatgaaagta	gttcggtgtt	aggtatccag	151920
tagttgacac	tggaagacag	ggaatgacat	gttaatattc	atagccagag	ggtggcccag	151980
gttttttcgt	acatgggaat	gaaattctta	tccaaataag	tagaaattat	gtgcgtaagc	152040
catttgttaa	gagcactgag	tatgtgcatc	tcgatccatc	taatgaataa	ccattatcac	152100
cagtttaaat	tattttcttt	aggcccagga	agagctagct	tggaagattg	ctaaaatgat	152160
agtcagtgac	attatgcagc	aggctcagta	tgatcaaccg	ttagagaaat	ctacaaaggt	152220
aaggatgact	tcgttttgtg	taaactaaaa	agtattattt	tccaggtgta	aaaataaaaa	152280
	gggtttcttt					152340
cataagcaac	tagaaaattg	acaaactaaa	tgaaacaact	gtttgcatat	attggacaat	152400
gggcaataca	gggaaaccat	ggaaaccaaa	cagageceag	tagtcttgct	gaacgaaaga	152460
gttaaatatc	aaagttcagg	ccaqqtqcaq	tggctcacgc	ctgtaatccc	agcactttgg	152520
gaggccaagg	cgggtgaatc	acttgaggtc	aggagttcaa	gaccagcctg	gccaacatgg	152580
tgaaaccctg	tcttagccgg	atataataac	aggcacctgt	aatcccaact	atttgggagg	152640
ctgagggagg	agaatcgctt	gaaccaggga	aacaaaaatt	acaataaacc	gagatcacac	152700
cactgcactc	cagcctgggc	gacgagcgaa	accccatttc	aaaaaaaaa	tcaaaqttca	152760
	tttgagtaga					152820
	gtagagatat					152880
	ggggtgaaaa					152940
gracatgege	gctggaacat	aggactaage	aggragaga	acceceage	tetageagg	153000
ctgaacgacc	ttactaatac	agggccaaga	adageteate	ttcacacaca	ttatgcctta	153060
gragagagac	ttagtaatac	acaayycacc	gggtagtgtt	ccacagagg	aagcaaattt	153120
ctactgaaga	taaattagtc	ccagagtaca	agcaccigaa	ccaagtttca	aagcaaaccc	153180
LLAAAGGGCC	aaattaccta	acaactycat	gucaaaacaa	aggictaacc	tecaecetes	153240
Laacacaaca	aaattcagca	cccacagig	Ladayttaga	acyccigacy	tacatores	153240
	tcatgcctgt					153360
gaggtcagga	gttcaagacc	ageetggeta	acatggtgca	accoegecee	cattadadat	153360
	agccaggcat					153420
ggcaggagaa	ttgcctgaac	ccaggaggtg	aaggttgcag	tgagccgaga	Legeaceact	T33480

gcactctggt	ctgggcaaaa	agagcaaaac	tcaggctcaa	aaaaaaaaa	gaatgtctga	153540
catcaatcac	aaattaccaa	gcatgacatg	aagttgacct	ataaccaqqa	qaaaactcaa	153600
totatacaa	cadacccada	tataaaaaa	atgatgaatt	tagcagacaa	agaccatcaa	153660
ataactattt	tasatattaa	asatatotto	aagtggccag	gtgcagtggc	tcatacctat	153720
actoccacce	ctttaaaaaa	ccaacataca	taggagttca	agaccagett	ggccaatatg	153780
aattttagta	ttetetaeta	2222522	aaaattagct	agacataata	gcaggtgcct	153840
gigaaacccc	ctctctacta	aaaacacaaa	cascastese	5550005505	gaggtggagg	153900
atagteedag	ccacacggga	ggccgaggca	caagaatcac	aggatage	3033033033	153960
ttgaggttgc	agtaageega	gattgtgcca	cttgtactcc	ageceggaea	traraaaaat	154020
actotgtoto	aaaaaaaaa	aaaaaaagc	taaagaaaac	aayaytataa	cgagaaaaac	154080
gcaaaatagt	tttaaaagaa	ccaaatggaa	tttcttaaaa	Ladadacac	cagaaacggg	154140
ggccgggcgt	ggtagctcac	gtctataatc	ccagcacttt	grgggggcrg	aggcaggcag	
atcacctgag	atcggtagtt	caaggccagc	ctgaccaaca	tggagaaacc	teatetetae	154200
taaaaataca	aaattagctg	ggcgtggtgg	cgcattgcct	gtaatcccag	ctacttggga	154260
ggctgaggca	ggagaattgc	ttgaacccgg	gaggcagagg	ttgcggtgag	ctgagattgc	154320
accagtgcac	tccagcttgg	gccacaagag	tgaaactccg	tctcaaaaaa	aaaacaaaaa	154380
aaaacagtag	actcgaagaa	ctagctgagt	ttttctttac	tttaggcagt	aagtgtgacc	154440
ttttqcaqqt	qactacttta	gttcctcatg	tcctcattag	tagatcagag	aaattegaca	154500
ccaaaacccc	aaaagaaaaa	ccccttctaa	tcctcattcc	atgattttat	gaatgcatga	154560
agtcctaggc	ctgcgaagga	atactcattc	tctttatcct	gtgttgatac	ctctctgctt	154620
caacctccaa	ctcgacattt	gcctatagga	tgtacttgga	cattcagcat	aaactacctc	154680
acaccattac	tgaattgctt	catgtgcaca	tgtcccatgc	cacaataccg	gggaccttgt	154740
cttccgtgat	atttgtccgc	agtgctgtga	ctacaggagg	gagtcagtga	atgtctgcat	154800
atatatettt	accatccctc	ttgaatatgc	tctagggtta	attcctagaa	gtagaattac	154860
tctattgaaa	attggcaata	tttttcattc	taatatctat	tgccaacatg	ggaaagcaag	154920
tctggatgcc	agteettgtt	atatqcccct	tgggtaagtt	acgtaacctc	tttaagcttc	154980
tottcactca	tattttaaca	aggaaaatta	caatatttta	cctcacaaaa	ttgtagtcag	155040
cttctaacta	tottaaacto	togtatatag	taaacactaa	atattaatat	ccatccttaa	155100
tttgtaataa	taggtcactt	gttagagaaa	tgcaccttac	cattttcttt	tcttttcttt	155160
trtcagttat	gactcaaaac	ttgagataaa	ggaaatctgc	ttgtgaaaaa	taagagaact	155220
tttttccctt	gattagatta	ttcaacacac	ccaatgaaaa	cagcactata	tttctgatct	155280
atcactatta	tttccaggag	agaatgggag	acaatcctag	acttccacca	taatqcaqtt	155340
acctotageca	ataattgatg	cacatgatgt	tcacacagtg	agagtettaa	agatacaaaa	155400
tagtattatt	tacattacta	gaaaattatt	agttttccaa	tggcaataac	ccatttatqa	155460
gagtatttta	acctactaga	atagacaggg	accacatcct	ctgggaagca	gataagcata	155520
gagegeeea	ttgatgcaca	ctcatagtag	taactcatcc	ctaatcagca	ttqtaaaqca	155580
gaaccgacac	ataatttact	ttatcettee	aaagcaggtg	agtcagcccc	accgagagcc	155640
aggegeeagag	geggeeegee	ataatactaa	cagcttcagc	ggaacagggt	gagagttaat	155700
tatacaatet	tetterese	gaggagaaa	tggaaggaaa	ctgacaagtc	atgggtcaag	155760
thteretere	tteeteette	ctctcatcc	agtatatagt	tttcacattt	taattcctcc	155820
tectagigac	cocccccc	anancette	tctcccctgc	taacacaacc	gratgaatet	155880
cottyagaty	caccatacce	atttgggggt	ttggaattct	ccactccact	tacttaactt	155940
ggtttactt	gagearragg	actiguete	agaaatagg	atgactccac	ttttacccct	156000
teeetteaga	atacatgtgg	adayadayad	agaaatagcg	ttttttt	taaataaggt	156060
graggeacerr	gaacaaagca	geteteeca	aattatactt	tagatagaga	attettteet	156120
gagcaggatg	actggggaga	gagaaacatt	tgactttgac	tratttagga	tactaacaca	156180
grgagergga	aagtgtgcag	tragtrager	ttcttctcct	cottagga	castcagaga	156240
ctcactcact	geacttetge	ccagttggct	tctgcatcgg	gattacatag	ccattageag	156300
gactgcccag	EEggEgagea	cactecattg	accacgtggc	geeagegeet	cettaatgea	156360
catgattgag	aggaaagaaa	gttctcttag	atgttactgc	ttttgctcag	actitiguada	156420
aaaaaaata	tatatatata	tgtataaata	tataattatt	aatcactttt	gcccccgaga	
aagtcttgaa	tgaacagaga	atttattcca	ttgcaatatt	tgattgtata	gaggcacact	156480
gtttcatcga	cagaagaagc	aaaaaggctt	tgtgtaagtt	tttggtacta	tgtaccacct	156540
			gtacttaaac			156600
			atttaaaatt			156660
gggcatttat	ggctcttctg	ttgaaatata	ttgatctttc	caaatatttt	catttgcttt	156720
ctaaaaaccc	agaacatgag	ccactactgg	actttgcctt	gtgtttgaag	tgtatggcat	156780
aaacccaagg	tttttattag	tcatctatgc	tgtgattaat	tcattttgtt	cttttaacaa	156840
aatatttcca	tccacttcac	attgcttcaa	tctttaacag	aaaagcaata	taaaggttat	156900
agaataaaat	gtggttttgg	gcaactcttg	ctgcctctgc	atgttttgga	ataacaattt	156960
ctacaagact	ctaggctgtt	taaactagtg	ctttcagtta	agataaattc	taatcatttc	157020
tttgtatata	cattttgtgc	ttctgagcta	gagatgccaa	gtagttgtaa	actgcttáta	157080
_			•			^

aagagaatag	cagcaaattt	gagactcggc	tactttttc	tgccccacct	gctttgagac	157140
acagaaggg	agtgtggccc	qaaattatta	gccagattta	atatttgatc	taaagtaggt	157200
cettetacte	attttaaagt	tagaatttga	ttcctccaac	attgagcacc	caccatgttc	157260
caggetetat	gcattgtgcc	cacaaaataa	gattccctgg	tggagttttt	atgggttcaa	157320
ataatcagtt	gaacaccctt	catctttatc	atattattaa	cattgacaca	aattgtttaa	157380
aaagaaaga	tattagagag	aaagtggtac	ctttqtaact	tgatgtgtct	tcatcattcg	157440
gtaagatttg	atgaaagtaa	aaagcaaatg	tcagccaaat	ccagtgaaca	gcaataaaac	157500
accasates	tttttataac	tttttctact	togatttcaa	cattcagtag	agcttttcga	157560
agggagtaat	ctttacacta	ctggaggttt	gactagttca	gtaggaattt	ggagggaag	157620
aatgtaagta	attataacaa	agtacaaact	tetttactat	tttatttaag	tactgagage	157680
gicaliciga	atcatcac	tgacctctct	ccagtgacag	totttaggta	cctgcctgac	157740
taagcacctg	acgaagcgac	ttctacacag	tgaccttttc	tetegeete	tectecetet	157800
ttcaggagtg	gggcccacgc	ggacctgggt	tgacctcctc	atccaracan	acccaagaca	157860
Egeceacaca	ccagtigatt	ggacccgggc	es egat gast	catacatata	attocaacac	157920
gttcttaatg	ttaagaattt	tggggccggg	cacggcggcc	ttaggcccgca	accedance	157980
tttgggaggc	cgagacaggc	ggatcacttg	aggccagggg	ctcgaggcca	tacagasasas	158040
catggtgaaa	ccctgtcttt	actaaaaata	caaaaactag	cigggcatgg	cggcgcacgc	158100
ctgtaatccc	agctacgtgg	gtggctgaga	caggggaatc	gettgaacet	ggaggcggag	158160
gttgtgcaat	gagccgagac	cgtgtcactg	cattccagcc	tgggtgacag	agggagaece	158220
tgtctccaaa	aataaaaata	agaaaaagaa	ttttgggcta	ggtgcagtgg	ctcacgcctg	
taattacagc	attttggaag	gcccaagatg	ggcagatcac	ttgaggacag	gagttcgaga	158280
ccagcctgga	caacatqqtq	aaactccatc	tctactaaaa	agacaaaagt	tagccagatg	158340
tggtgatggg	cacctataat	cctagctcct	cgggaggctg	gggcaggaga	atcacttgaa	158400
cccaggaage	agagattgca	gtgagccaag	atcacatctc	tgcactccag	cctgggcaac	158460
agagcaagac	tctatctcaa	aaaaaaaaqa	atttggccag	gcgcagtggt	tcacgcctgt	158520
aatcccagca	ctttqqqaqq	ccaaqqcaqq	cagatcacga	ggtcaggaga	tcgagattgt	158580
cctggctaac	atggtgaaac	cctgtctcta	ctaaaaatac	aaaacattag	ccgggtgtgg	158640
tagtaggcac	ctgtagtccc	agctactagg	qaqqctgagg	cagaggaagg	atgtgaaccc	158700
aggagggga	acttacaata	agccaagatc	gtgccactgc	actacagtct	gggcgacaga	158760
graagactcc	otctcaaaaa	aaaaaagaat	tttqqccqqq	tgcggtggca	catgcctgta	158820
gtcccagcac	tttgggagac	caaagtgggc	ggattacctg	aggtcaggag	ttcaagacca	158880
gtccggccaa	tatoocgaaa	ccctgtctct	tactaaaaaa	aatacaaaaa	ttagccaggt	158940
ataataacaa	gcacctgggg	aggctgaggc	aqqqaqaaat	gcttgaaccg	gggaggcaga	159000
gattacagta	agccaagatc	gtgccactgc	actccagage	aagactcttt	ctcaaaaaaa	159060
99229222	aattttgcat	ggggaaggag	agatactott	caccatctgg	aatggtgctt	159120
ggatgfggg	cttacaaaat	caggagccag	cactgcatgg	acaaacagaa	gcatgtgggc	159180
ctgagatage	aggtaccttg	ataaccctga	agacatcctt	ggtttctgca	tctattcctg	159240
cateetteea	tropactaca	ttaatctgtc	agttatcctt	ataatgattt	ttgattttt	159300
ttttttaaa	tagaatttag	ctcttgttgc	ccaggctgga	gtgcaatggc	acqatctcgg	159360
ctcacacac	cctccacctc	ccaggttcaa	grgattctgc	tacctcaacc	tootgagtaa	159420
ctcaccacaa	aggeatocce	caccacacct	ggctaatttt	gtattttag	tagagacggg	159480
cigggattat	attesteses	ctggtctcga	acteceaace	traggtgatc	accetatete	159540
gettectat	actactaca	ttacaggcgt	aagccatggt	accoggicta	ttttttgatt	159600
ggeeteccaa	agtgttggga	gagtttttt	aagccacggc	aaggagtttg	gctaaaatac	159660
ttttgaaacc	agecegaage	taatgattat	gtatteteag	catototoca	aagtactgct	159720
tgccatactg	ccctaatgcc	tctttagtaa	acttoactta	agtcgtcatg	totattotot	159780
gatttetgga	gaataattt	atggagctaa	22626666	trattettat	aacaagcagt	159840
caaaatggta	teetaaeeta	atggagetaa	aagacacccc	atgatcacca	ttagaactca	159900
tactgaggcc	caggaagggg	agaagtccct	ggettgtgag	acgaccacca	gaagattata	159960
aacctaaacc	agtgeettt	catgettete	agateettee	tttttaatt	gaagactaca	160020
accecttta	gcaattgtaa	taaacccaga	aacagaaagc	atacataca	attetatata	160080
agaagtttgg	cgggagagat	aatttttaca	aaatttgtaa	acaccigcca	acccacaca	160140
ctaggcaagg	tetetggeet	tgtaaaaccc	ctcaaggtta	caactttggt	ggedeacace	160200
aatagttacc	cactgaggcc	ctctccgggt	gaacattgag	cactagagga	agececters	160260
cttgggcagg	actgggcgtg	gtgcagagta	ggagcggtga	tactgtggat	cctgggcagg	
tggagatggc	cagtgatgtc	caataaagga	cactggaggg	agcagtgtga	gtaaaggccc	160320
tgagggcatt	catattcaga	gagggttgct	gcccactggc	ttgcttggca	cacaggagag	160380
taggtattcc	tgccttagta	actttatqta	aacaagtatt	tcctcagtct	gttcctctca	160440
aactgcctgc	tctqqcacat	tcagaatgtc	acagaactca	cctggatgca	ttcagcccct	160500
tacctaaaqq	tgacagtgca	tctccttccc	caccccaccc	ctcataccac	tgaagcacct	160560
gtcagactgg	cccaqtctqt	gggcaaggag	cctagagagg	gcttagtttc	agcttgaaag	160620
gagctgggat	ttaccaagaa	gcaaatgaga	gacgaggatt	gcaacaactg	tgccatttcc	160680
	-			_		_

WO 01/27857

```
ccagcttcag ctgactcctg tatattgact gtgccttcag actcatccgt aagtgacccc
                                                                                  160800
aggetggeet eteccacate acagtaagaa ttecacacae catacaactt ggaaagagge
tecagetgaa ggaageeeca cacttette aagtttttet tagtettete ttettggcaa
                                                                                  160860
agagtacett ttgtttette taattatgta actattggtt tagtaaatat teacceatte
                                                                                  160920
                                                                                  160980
agreacetty taagtggeag geactgritta cagggacaca ggaaggaata aaaacttgca
ggcaccttgg agcttgcatt ctattgaaga ggtaatggaa gttgggatag cagctaaact
                                                                                  161040
atgetggtat tggecaggeg cagtggetea cacetgtaat cecageaett tggaggecaa
                                                                                  161100
ggtgggcaga tcatgaagtc aggagatcga gaccatcctg gctaacatgg tgaaaccccg tctctactaa aagtaaaaaa aaaaattagc caggtgtggt ggcgggcgcc tgtagtccca
                                                                                  161160
                                                                                  161220
                                                                                  161280
gctacttggg aggctgaggc aggagaatgg tgtgaaccca ggaggcgaag attgcagtga
gccgagatgg caccactgca ctccagcctg ggtgacagag cgagactctg tctcagaaaa aaaaaatatg ctggtagttt tgattcaaga tggcctttgg agcccatgat ttaggtctcg
                                                                                  161340
                                                                                  161400
tacccaccaa ggtctactgg aaaacatcag gctctcctgc tatagaccca tagggagagc
                                                                                  161460
                                                                                  161520
tgcagccgag agggggagct gaagagaagt gccccttctg tgtcctgtca gcctcatcct
                                                                                  161580
teegeaagga ceagttgetg tgeeacteea tteacttget geaagactgg aggtttttee
traggtgttg agracetggt ttaraagatg traggatett gatgertgag accatraagg caagtetetg aaragggett accttagagt aaggettaga agaggergta aagtragtet
                                                                                  161640
                                                                                  161700
cagctccgtg gctctgcaga gctttgggac atgtgaattc ttaaaaacaa gactattgta cagttactat atgcatgcag tataaaatta taaccttgga aaatcctagc tagctgttga gctaattcca taaagtaatc agctcctgag ttctgcagtg gtaataataa tcagcataaat gagtaaacac tgtgtgtgc aggcagcgtc tcatttgatc cttgtgataa tcttgtaagt
                                                                                  161760
                                                                                  161820
                                                                                  161880
                                                                                  161940
                                                                                  162000
162025
tgttgcccag gctagtcttg aattc
<210> 36
<211> 162025
<212> DNA
<213> Homo Sapien
<220>
<221> mutation
<222> 156,277
<223> Nucleotide Base Change: T to C
gaatteetat tteaaaagaa acaaatggge caagtatggt ggeteatace tgtaateeca
gcactttggg aggccgaggt gagtgggtca cttgaggtca ggagttccag gccagtctgg
                                                                                       120
ccaacatggt gaaacactgt ctctactaaa aatacaaaaa ttagccgggc gtggtggcgg
                                                                                       180
                                                                                       240
gcacctgtaa toccagctac toaggagget gaggoaggag aattgcttga acctgggaga
tggaggttgc agtgagccga gatcgcgcca ctgctctcca gcctgggtgg cagagtgaga ctctgtctca aaaagaaaca aagaaataaa tgaaacaatt ttgttcacat atatttcaca
                                                                                       300
                                                                                       360
aattigaaat gitaaaggta tiatggicac tgatatccig titcattcit tatataatca
                                                                                       420
                                                                                       480
ttaagtttga aatgtatact tgcactacta acacagtagt taatcttagt cctacaagtt
actgctttta cacaatatat tttcgtaata tgtatgcact ggtgtttatg tacgtgttta
                                                                                       540
tgtttatate tgttaaaatt ageagtttee atettttet attttgtace ateacateag
                                                                                       600
ttcagaagga ttgacagagc aaaatgattt gatgaagtat aaaagtcaca tggtgagtgg cataaataca actctgaaca attaggaggc tcactattga ctggaactaa actgcaagcc
                                                                                       660
                                                                                       720
agaaagacac atatcctata tgtcaagaga tgtaccaccc aggcagttaa agaagggaag
                                                                                       780
tacacataga aagcacaatg gtgaataatt aaaaaaattgg aatttatcag acactggatt
                                                                                       840
catttgctcc taaagtcaga gtcctctatt gttttttgt ttttgtgggt ttcttttaa
                                                                                       900
attttttat ttttgtaga gtcggagtct cactgtgtta cccgggctgg tctagaactc ctggcctcaa acaaacctcc tgcctcagct tcccaaagca ttgggattac agacatgagc
                                                                                       960
                                                                                      1020
                                                                                      1080
cactgagccc agcccagacg ctttagcatt tatgaagctt ctgaaatagt tgtagaaacc
                                                                                      1140
gcataagctt tccatgtcac tttcaaagtt tgatggtctc tttagtaaac caaccaagtt
attecteaag ggeaaaataa cattteteag tgeaaaactg atgeaettea ttaccaaaag
                                                                                      1200
gaaaagacca caactataga ggcgtcattg aaagctgcac tcttcagagg ccaaaaaaaa
                                                                                      1260
aggtacaaac acatactaat ggaacattet ttagaagage eccaaagtta atgataaaca
                                                                                      1320
                                                                                      1380
ttttcatcaa agagaaaaga gaacaaggtg ttagcaaatt cctctatcaa ataacactaa
acatcaagga acatcaatgg catgccatgt ggaagaggaa gtgctagctc atgtacaaac cagtagataa tttcaacttg ctgccgaatg aaacctcttt gcaaggtatg aatcagcact
                                                                                      1440
                                                                                      1500
```

						1550
tctcatqttt	qttttgcttt	gttttgtttt	gtttttagag	acaggccctt	gctctgtcac	1560
acadoctoda	atacaataac	acqatcaqaq	ctcactgcaa	cctgaaactc	ctgggctcaa	1620
acaggoogge	taccttagcc	tcccaagtag	ctaggactac	aggcccacca	tgcccagcta	1680
gggattette	ttttttatata	agatgggatc	tractagrac	ctttcatgtt	tgatgttcat	1740
attittaaa		tatagaggaca	catatagaaa	atctaaagtg	aaddaddacc	1800
atacaacgac	caaggtacaa	tgtggaaaag	99000000	acttaataa	adgaggaco	1860
agaaagaaaa	ggggttgcta	catagagtag	aagaageege	accidatgcc	agectacaac	1920
actgctgttt	tcctcagagc	agagttgatg	atctaaatca	ggggtcccca	accccaget	
catageetgt	taggaaccgg	gccacacagc	aggaggtgag	caataggcaa	gcgagcatta	1980
ccacctgggc	ttcacctccc	gtcagatcag	tgatgtcatt	agattctcat	aggaccatga	2040
accetattqt	gaactgagca	tacaaqqqat	qtaggttttc	cgctctttat	gagactctaa	2100
taccagaaga	totateacta	tcttccatca	ccctqaqatq	ggaacatcta	gttgcaggaa	2160
2202324454	gaacteccat	tgattctata	ttacagtgag	ttotatcatt	atttcattct	2220
atcatcccca	gggctccac	tagaaataaa	ggcacaatag	accadacata	ataactcaca	2280
acactacaac	graaraataa	ggaggccaag	ggaaaaaaa	cacgaggtca	ggagatcgag	2340
cetgtaatee	cagcactteg	ggaggccaag	ctactaaaaa	++c======	aattagccgg	2400
accatcctgg	ctaaaacggt	gaaaccccgt	t		accageogg	2460
gtgtggtggt	gggcacctgt	agtcccagct	actcgagagg	ctgaggtagg	agaacygcgc	2520
gaacctggga	ggcagagctt	gaggtaagcc	gagatcacgc	cactgcactc	cagcocgggc	2580
gacagagcga	tactctgtct	caaaaaaaaa	aaaaaaaaa	aaagaaataa	agtgaacaat	
aaatgtaatg	tooctoaatc	attccaaaac	aatcccccca	ccccagttca	cggaaaaatt	2640
ctcccacaaa	accagtccct	agtaccaaaa	aggttgggga	ccgctaatct	aaataatcta	2700
atcttcattc	aatqctaaaa	aatqaataaa	ctttttttta	aatacacggt	ctcactttgt	2760
tacccagact	ggagtacggt	ggcatgatca	cagctcactg	tagcctcaat	cacccaggcc	2820
ccaccoatcc	trocacetaa	acttcctgag	tagctgggac	tacaggcacg	caccaccatg	2880
ccagcgaccc	ttttaaattt	tttatagaga	tagagatete	accatattac	ccagactggt	2940
cccagccaac	cccaaaccc	gatecteect	caaactcctq	gactcaagtg	atcctccttc	3000
CLCaaacccc	gggcccaagc	ggattacaag	catgagggag	tatacccago	togataaaca	3060
ettggeetee	caaagtgctg	ggaticacaag	acgageeae	acatocacta	tagacagtaa	3120
ttttaagtcg	cactacagte	atggacaatc	aggererea	acacgcagca	aggueagega	3180
gtcccagggt	ctgcttttcc	atactgaaat	acatgtgata	Ctaaggagaa	aggigetege	3240
aaggatattt	aaaatgaaga	atatttaaaa	tgaggaaaaa	actgtttctt	catgacttg	3300
ataaggctga	taaagaccat	ttctgtgatc	tcaggtgatt	cactcaagta	gtatatttca	
otaatcatta	tetagaacag	cctgaatctt	aaccaaaata	ccatgatttt	LLaatguigu	3360
tatgatacct	taataatata	accaaactgc	aatgtaggca	gctaaatctc	cacgagtttg	3420
acttccccga	gagttgagag	ttttcttcac	aaattaaaga	aatatattt	ttgatacatg	2400
attogcatat	ttaaaaacta	cactgaaatg	ctqcaaaatg	atataaagaa	acattttcca	3540
caatcáááto	caatcaaaga	gtagattagg	aatctactca	ccattatcaa	ctaaatagaa	3600
acacttocac	tagatataat	ggeteacate	totaatctca	gcactttggg	aggccaaggc	3660
acacceggae	cttgaggca	ggagctcaag	accagcetga	gcaacatagc	aaaactctgt	3720
aggregating	222222222	attaaccagg	catogtogca	gatgettgta	atcccagcta	3780
CLCLacada	tanaataaaa	ggactgcttg	saccesacea	atcaagactg	cagtgagccg	3840
ctctggaagc	Lyaagtagga	ggactgcttg	ageceaggag	ctatatas	aacaaaaaca	3900
tggtcatgct	gegecaeage	ctgagtgaca	gagagagace	ttattatta	tttatttta	3960
aacaaaaaac	acttaacctt	cctgtttttt	gorgregerg	cetestests	actosctoca	4020
gatggagtct	cactctgttg	cccaggctgg	agrgcagrgg	egigaliting	gettaetgea	4080
agctctgcct	cccgggttca	cgccattctc	ctgcctcagc	ctcccgagta	gergggaera	4140
taggcgcccg	ccaccacgcc	cggctacttt	tttgcatttt	tagtagagat	ggggtttcac	
catattaacc	aggatggtct	tgatctcctg	acctcqtqat	ccacctgcct	cggcctccca	4200
aagtgctggg	attacaggca	tgagccaccg	cacccqqcca	acctttctgt	tttttagttt	4260
gatatgcttg	ttaactcagc	agctgaaaga	atqctgaaag	tggccttcag	taaaaaaatt	4320
tcactagaat	ctctacatcc	atatttaatc	tgaatgcata	tccagattga	tcagttagag	4380
cassascact	catcatcatt	cctgatgacc	tctaattctq	qtttcggctt	tetattteaa	4440
tagaaacaca	ataaggaaag	aaatggaagg	gctctggaaa	tttatcctaa	gctatagata	4500
cygaaacaga	tcaccaacaa	taagatetet	cctataaata	taaaacaagt	ataattaatt	4560
ctattaaaga	tttttatat	tcagaggatt	ttatttcaag	ataaaacata	acttctaccc	4620
ctttaattat	++000000CC	tagaaaaagt	attttcctc	atcttatcct	tcaaagaggt	4680
atactattga	LLCCaaaggc	Lagadadagt		atcocacct	atactccaca	4740
cacagcaatg	caaacatcta	taaaatgcct	Ligitatel	tatattaaa	aacctctatc	4800
aatcattgaa	aatgcttttc	cattttaagc	ccaggcgagg	tgttttagga	testetate	4860
acaacttact	ctatttattg	ggaggtaaac	tcccagactc	receagggre	Legigiacig	
atctcatttt	ttaggcttcc	taatcccttg	aagcacaatc	gaaaaagccc	tggatetett	4920
ttctgcacat	atcatcocoo	aattcattcg	gcttccagca	agetgacaet	ccatgataca	4980
agcagectica	cccttctccq	gacgccagtc	cttgctgcgg	ttagctagga	tgaggggttt	5040
actagacttc	agtgcaggct	tctgcgggtt	cccaagccgc	accaggtggc	ctcacaggct	5100
J = - J J						

ggatgtcacc attgcacact gagctcctgg caggctgtac caatttttta attatttaat 5160 atttattttt aaaattatgg tgaatatttt ggtattctgc tctaaaatag gcccataaat 5220 gcacagcaga tatctettgg aacccacage tttccactgg aagaactaag tatttttett 5280 ttaaagatge tactaagtet etgaaaagte cagateetet acetetttee ateccaaact 5340 aagacttgga atttatgaga gatctagcta acagaaatcc cagacacatc attggttctt 5400 cccagagtgc agtcctccta aagaggctca gccctaagca ggcccctgca ccaggagggt 5460 gggtctgaga cccacatagc acttcccaag gtgcatgctc cagagaggca ctgaaacagc tgagcacaag cctgcaagcc tggagaactc tcacagtcag aacggagggg gcccagtggg 5520 5580 actaacataa agagaaaagg gaacacagag aaatggatgg caccaacaac cagcaaagcc 5640 ttcatggcca atgaaagcat cagtgacggg gccagaaccc tcatccccaa agactcttca 5700 ctgcctttag tgaaaaacaa tggctagaga gtgaagttat gatcatgtat agagaggtaa 5760 5820 agttacattt ttatattctg actctgctaa tgtgaaattc cctatctgct agactaaaag titcagacac cotgitcaaa tatoccatta gitgotagag acttaaaatg aacagaacgo 5880 acattotcag gatgactatt accaaaaaat caaaagacag caagtattog tgaggatgta 5940 gagaaactgg aacttttgtg cactgtttat gagaatgtaa aatggagcag ctgctgtgga 6000 6060 aaagagtatg caggtteete aaagagtaaa accaagatgt ggaaacaact aaatgeeeat 6120 cagtggatga aggggtagac aatatgtggt atatacatac catggagtac tattcagcct ctaaaaaaa aaaaggaaat tctataacat gcaacagcat ggatgaatct tgaggacatt ttgctaatga aataaggcag tcatagaaag acaaatactg cacgactcca cttatatgag 6180 6240 ataccaaaaa tagacaaatt catagaatca aagagtacaa tggaggttac ctggagctgc 6300 6360 agggcgggaa acgaggagtt actaatcaac gaacataacg ttgcagttaa gtaagatgaa taagetetea agateagetg tacaacaetg tacetagagt caacaataat gtattgtaca 6420 cttaaaaatt tgttaagggt agattaacaa atgtagtaga tccacaaatg tggttaagtg 6480 ttottaccac agtaaaataa aaaaagaata tcaagcccag gagttcgaga ctagcctggg taacatggtg aaaccctgtc tctacagaaa atacaaaaat tagccagctg tggaggtgca 6540 6600 ctcctaggga ggctgaggtg ggaggcttgc ttgagcccag gaggtcaagg ctgcagtgag ccatgattgc accactgtac tccagcccag atgacagagc aagacaccac ccccccaaa 6660 6720 aaaagaaaaa gaatatcaaa cattttaaaa gatcagatac gcaagaacaa caacaaaaaa 6780 gagatgaaca gagcatcgac cctcatctag tgggattctt ggtctaactg aaaaacagac 6840 attgagagac aaacaatgac agtgatgtga tcacagcaat tacacaggta tcccctgggg 6900 actgcagaag aaaggaggaa tgcctaactt tcagaaaata gagaaagcgt caaacagttg gtgaaagcct tccaaaacta gagagaactg cacacaccaa atcacagaaa gaagaaaagc 6960 7020 cgtgggagat tctgggaccc accggctatt tttgatggct gaacaccctg ctgcaggaga 7080 gacaggaget ggaaageatg gtgggatgaa acetcaaaca getttgeetg cattgettaa gatgaetggg ettgattaae tetagteaat ggggacaatt caatcaaaga agaaagatge 7140 7200 7260 tcaaattcac attttagaat gattttttat ggcagtatgg ggaatagatt aaaagagagt 7320 gaagetggag geaagaaact tgttaagagg caactgaaac agtetagatg ataaataata aactgacaga gtgactagaa aaatcagaac aggctgaatc aacagatacc tagatgaaaa taacaggact tgatcaccag ttgtatcttg gagaggaagg agttgtttcc ttgctttccc 7380 7440 tacgactggg aatacggaag gtitgccgtg igtaliggii alatactggt gigtagccaa 7500 7560 tcactgacaa ccatttagca gcttaaaaca caaaggctta tctcccagtt tctgtgggcc aggaatctaa gataggetta getggetggt tetggeteag agttteteaa gaggttgeaa 7620 tcaagatgtc agctggggtt gcatcatctg aaggctcaac tgggggccgga gggtccactt 7680 ccaaggagtt cactcacctg cctgacaagg cagtgctggt tgttggcagg agatctcaat 7740 7800 tcattgccaa gtgagcctct ctatagcatt gctggaacat cctccccatc tggcagttgg 7860 cttctctcag catgagtgat ctgagagaga gagcaaggag gaagccacag tgttcttcct actectacte ctaacactat ggacetacte ctaacactet cacttetgee ttattecatt 7920 agttagaaag ggaactaage tecacetett gaaataagaa gtgteaaaga atttgtggat atatttaaaa atcateacae tgtggaagtg gatagggggt teaattaatg etgaaettga 7980 8040 aatgeetgag acatteaaat gteeaacagg caatgaacat acceatagat ggteatgaet 8100 8160 ttaqcaagaa tagaggaaga tcacagaatt aaggaggaat tgaaaggtaa aagaagtgga 8220 gtcagattcc ccctgaaaag tgagccatga aaggaacttt aactattgag ttagaggtca gagtaggaaa tttcggtgga attcttttt aaagaaagga accatataag catgttttga 8280 8340 ggtagaggga gaataaatca gtagacaggg agaggtaaaa aacataaatg ataggggata 8400 gttgacaaag gtcttggcag aatcccttac ccattgactt ggggccaaga gagggacact tetttgtttg agggataagg aaaataagaa agaatgggtg etatttagtg tggteetgte tetagggeaa acgeataggt aacaaactgt gtgtgttagg aatatagatg tgaceteaca 8460 8520 ttgagattet cacctcaaat ccattttgtt gttacctgta cettectace ttetetttt 8580 8640 gctacatgca gactgctgtt ttgtcttcct ggcctgttcc aggtttcagc attctggcat 8700 atotgotace etgiteceaa acctetetag agreeatget cetteettgg atagtgittg

			tasattagge	ctgagaacct	cctctatoga	8760
attgggccac	gtatctaaga	agtgatgcct	teagtragge	atgtgagtgc	teccaaceta	8820
aatctccatc	agtgaccctg	acagacttgg	cateteggag	atgecactge	agaratet	8880
tggtctagga	gaatctcagc	ctgggcctct	agtagtatgg	acaaggegee	adgetteete	5 8940
tgaaccagag	tctgtcatat	tcctcaatgt	gggacagaca	atatagegge	tttcagagg	9000
tttctgagct	agaactctgg	tttttggtct	agattettig	atguatgace	attagagge	9060
attaaaattt	gttctaatac	aatgttcaat	acaaatgtag		greaggacer	9120
caacaaaaca	tgaccaactg	tagatgaaca	ttaaactatg	acaacccatg	gaaatgaata	9180
cartaatacc	tacaattccc	ccattttagc	agtcactatq	gracating	ycacaaatyy	9240
ctatttaagg	gtgcttttgt	taaaacctac	catcttacta	ggcacatgat	accgaaacca	9300
atgaaataat	ggagaaactt	cttaaaaact	tttaatgaat	aaagtgatga	agegataata	9360
ttttagctgc	tatttataaa	gtgactatta	caggicaaac	attettetag	ggttttttg	9420
rtgaagttgt	cacatttaat	ccttaataac	ccactatgag	tcaggtattc	ffefefeee	
tttggacagt	tagagaaata	aggatcagag	aggttaggta	attigctcag	ggccacacaa	9480
cctacatata	gaaaatctga	gatttgtaca	ggaacgtatc.	aaactctgaa	greeargerr	9540
Ctattttccc	atgetgeett	totaataaaa	ggtaactaat	gctactggat	getgeeeca	9600
aagtgagtca	ctttcacccc	accctacttq	attttctcca	taaaactaat	cacatectga	9660
caacttattt	attoctoatc	tececcaeta	gattataaac	tcaataaaag	Caagateett	9720
atctactaaa	tatcagtacc	taaaacgctg	tctagcacag	agcaagtaat	taatattigt	9780
tgaatgaaga	aataaaggaa	aaaaattcaa	aggaagaaaa	agccctaaaa	cagatgitta	9840
cctasacata	cattttaaaa	gaaagcatat	aacaaattca	ggacagaatt	taaatttgat	9900
tttttaaaga	aataaccaaq	tqctaqctqq	gcacagtggc	teacacetgt	aateetagea	9960
ctctaaaaaa	ccaaaacaaa	cagarcacti	daddicaada	quicaaqacc	agectiggeta	10020
acatootoaa	acctototot	actaaaaata	cagaaattat	ccaggcatgg	tggcaggtcc	10080
ctotaacccc	agctactcag	gaggetgagt	caggagaatt	qcttqaaccc	aggaggcaga	10140
aattacaata	ggccaagatt	gcaccactgc	actccaqcct	gagtaacaaa	gcaagactct	10200
ggctgcagog	gaaggaaaga	aagaaggaaa	gaaggaaaga	aggaaagaag	gaaagaagga	10260
3000300330	gaaagaaaga	aagaaagaaa	gaaagaaaga	aagaaagaaa	gaaagaaaga	10320
2202220222	gadagadaga	gaaagaaaga	accaagtgct	tatttgggac	ctactatgct	10380
aagaaagaaa	atguauguuu	ttttcagtaa	agcagttagc	aaacttgcaa	gatcataaca	10440
acgettetet	acgeacgeea	ctctaaaatt	gtgctttaag	aagttcctct	ttaccaqctc	10500
acaaacacac	tagttttgta	agagttacta	graacttttt	ccctggagaa	tatccacage	10560
acguatguat	agettetta	ggatgcttac	taacatgaag	ttatcaaatg	tgagcctaag	10620
tagectatet	testettast	atactccaga	acaaaaacca	tectactote	ctctgacaat	10680
Ligggerage	anttonttt	ccacattacc	aaddadccad	ggtaggagaa	tatagaaaga	10740
	adictatit	ctttcagcaa	aaggagccag	aagtaggtaa	ctaaacacat	10800
ccacccaaga	tasstagge	attgtgctca	gaagaatgat	ctacaacatc	ttactgtgaa	10860
gccccaacaa	nantattora	ataagacttc	tctccaaaat	gattttattg	aatttgcatt	10920
ggaactactg	ttttaagget	aaattttaaa	aggtttgata	ttggtacatg	aatagacaaa	10980
ttaaaaaata	etteaageet	aattaggttc	aggeetgaea	aggaatttaa	tatacgataa	11040
cagacatgga	ctagaccaag	aactaggtte	tattaaaaca	aggaactaaa	catcaggaaa	11100
acctagtatt	ccaaaggaac	caacaaatgg	202003000	taactgaagg	agcaccaaat	11160
aacacagttg	ggcaccctac	cttactccta	acaccaggag	tatacacaac	atgcaggttt	11220
atttattat	tttaattata	gttttaagtt	ctagggtacg	anatattaa	220222222	11280
attacatagg	tatacatgtg	ccatgttggt	gaggagcacc	ttaggaggg	aagaaaaaaa	11340
ttggccaggg	geggtggete	acacctgtaa	ceccageace	atgggaggee	cccatctcta	11400
gatcacctga	ggtcgggagt	tcgagaccag	cetgageaac	atggagaaac	ctacttccca	11460
ctaaaaatac	aaaattagcc	aggcatggtg	geacatgeet	graaccccag	ctacttggga	11520
ggctgaggca	ggagaatagc	tttaatctgg	gaggcacagg	ctgcggtgag	ccgagacacc	11580
gcactccagc	ctgggcaaca	agagcaaaac	ttcaactcaa	aaaaactaac	adaladalad	11640
aaataaagaa	agaaaagaaa	aaaatgaaaa	tagtataatt	agcagaagaa	aacaccgtag	11700
aatcctcgga	ctcttaggat	ggggaatgcc	tataatataa	aaaccccgaa	gitataaaag	11760
agaaaatcac	ctacatacaa	accaaatctt	tctacatgcc	taaaacatag	Cacaaacaca	11820
gctaaataat	catagctgaa	tgaactggga	aaacaaaact	tgactcatat	ccagacagag	11880
ttaattttcc	tacacataaa	gagtacctat	ataaacccaa	caaaaaaacc	accactaacc	
caaaataaaa	atgtgacagg	taatqaacaq	qtaqttcaca	gagaatacaa	atggctcttc	11940 12000
ggcacataag	atgctcagac	tgacttttac	ttatttattt	tttgagagac	agggtctcac	
gatgttqccc	aggttaggct	caaactcctg	ggctcaaatg	atagtaccag	gactacaggt	12060
atacccacc	gcacctggct	cctcaaccac	ctgtattaac	aggaaatgca	aaataaaact	12120
trcaaatcta	ttttacctat	taqaatqqca	aaaatttqaa	aaacttcaaa	catcatcatg	12180
ttootoagaa	tataaqqaqa	ctqqcactct	cattttttgc	tgatagcata	tatatactga	12240
tggcttctat	ggaaagcaat	ctggcagcgt	ctatcaaatg	tacaagtgca	tatatccttt	12300
	_					_

				atacttccta	agactagaa	12360
gacaaagcaa	ttccactcta	ggaatgtgtt	Clatategett	gtgcttcctg	9990099944	12420
ctgggagcta	agggacaggg	gcagaagata	atcttcttt	cectecttee	ccgccaaaca	
tattaaattt	tatatactot	aatatattat	ttttcacaaa	agataatttt	Laagegatat	12480
gtctgggaat	tttttttt	cttttctqaq	acagggtctc	actctgtcat	ccaggctgga	12540
atgccatggt	atgatctcag	ctgactgcag	cctcgacctc	ctgggttcaa	geaateetee	12600
acgetacgge	tectgagtag	ctgggactac	aggcacgtgc	catcatgcta	atttttgtat	12660
caccicages	teactatet	accesagets	atotcasact	cctaggctca	agcaatccac	12720
atacagggtc	ccaccacycc	geceaggeea	acgeotaace	caccococct	ggccctggga	12780
ccacctcagg	ctccaaagtg	cegggattac	aggegegage	caccgcgcct	9900009990	12840
attcttacaa	aagaaaaat	atctactctc	CCCLLCLALL	aaagtcaaaa	cagagaagga	12900
aattcaacct	ataatgaaag	tagagaaggg	cctcaaccct	gagcaacaaa	cacaaaggct	12960
atttctgaga	caggaatttg	ctgaacaaaa	tcgagggaag	atgacaagaa	teaagaetea	
cttctcaact	aaacacaata	gctcacacct	gtaatcccag	cactttggga	ggccgaggcg	13020
gacagatcac	gaggtcagga	gattgagacc	atactggcta	acacagtgaa	acceagicic	13080
tactaaaaat	acaaaaaatt	agccgggcgt	ggtggcaggt	gcctgtagtc	ccagctactt	13140
gggaagetga	ggcaggagaa	taacataaac	ccaqqaaqcq	gagcttgcag	tgagccgaga	13200
tracgreact	gcactccage	ctagataaca	gagcaagact	ctgtctcaaa	aaaaaaaaa	13260
aagactcatt	tetetagate	ttgagccgta	ttcaaattta	tctcagctta	gtgagaggtt	13320
aagaccaaca	atateettee	ctatagaccc	toctccttac	tgaaggaagg	taacqqatqa	13380
aaagcaagga	ccaatccaca	aaaccactaa	caccattato	tgatgaacat	tacgtgaaga	13440
gccaaggaca	ctaatggaga	attactacaa	aaatcaataa	aagcggacat	tcatttaggg	13500
agggtaagaa	grgaagrgga	attgctgaag	ttaaaaaaa	attagagacat	asaacaaaaa	13560
aaatggaata	taggaaatcc	acaaaagcga	Luadaaagat	gttagaggct	est et et set	13620
gaccacaggg	tcaggagatc	gagaccatcc	tggctaacac	ggtgaaaccc	caccccacc	13680
aaaaatacaa	aaaattagcc	aggcgtggtg	gcaggcacct	gtagtcccaa	etaeteggga	13740
gactgaggca	ggagaatggc	atgaacctgg	gagacggagc	ttgcagtgag	ccgagatcac	
accactacac	tocagostag	gtgacagagt	gagactccat	ctcaaaaaaa	aaagttagat	13800
acqaqaqata	aagatccaac	agacacacaa	ctqctaattc	tgaacagaac	aaaacaaacg	13860
gcacaggaaa	agaaaattta	agatataaca	ccqqaaaact	ttcctgaaat	tgagtaactg	13920
aatctatage	ttgaaagggt	ttagcatatg	ccaagaaaaa	tcagtagagt	ccaaccagca	13980
caagacacat	ctagcaaggc	taataattet	accaacacag	agaaagaagt	gggtgaccca	14040
taatgcggaa	aaaggcagac	catctgcagt	cttctccaga	acactggagt	ctgaagacaa	14100
aadaatooto	cctactgage	cagaagggag	agaaagtgac	ccaacacatc	tttaccaagt	14160
tagaatgttg	cccactgage	aaaaaaaaaa	asaccatos	aagacatgaa	agaacacaag	14220
Lagaatgtca	atanagasa	acaagecatte	tcatactcaa	gaatccttaa	gaaaaatgta	14280
Caltiacaac	acgaaagaac	acaugeacee	tacttaatot	gctcattaat	gggaacttca	14340
gtcctaatcc	ageceaetga	aagttaaatg	nacetatete	geceacttt	cctccaatac	14400
tagetteaaa	teagtetggt	eecatetace	aacaccccc	gcccggcttt	atageagag	14460
tcagcacctt	tecetectee	cagtettgte	ccctggagtc	tgctctcagc	acagcagage	14520
gaccacatca	acacccaagt	cagageeete	cagtgcgcac	tggtctacaa	agecetteee	14580
accccccacc	ccacgtgccc	tccggatcct	tgtgacgtgt	ctcctgcata	ccctagcagc	
cctggcctcc	tcactgcccc	tcctgtacat	caggaaggcg	actccttgag	tattggatat	14640
agecacetec	tecacetgea	gtgagttaac	tecettacet	actctaggtc	attqctcaaa	14700
totcagcate	tcaatggggc	cctccctgac	taccctattt	aaattctaca	tactcccctt	14760
gaccccatgg	acctcactca	ccctattcca	cttttattct	tacaatttag	cacttgttct	14820
cttctaacgt	attctaagac	ttactcattt	attacattqt	ttqccacccc	ctctagtaca	14880
taaactccag	agggggaggg	atttctqtct	atttattcat	ttctttatcc	ctaggacata	14940
gaacagggca	tagttcagag	tattcaatqt	tatcaatgaa	tgaactagca	gtagtaccag	15000
ttccagttag	gcacagaatt	aaatctaaat	agaattaaat	ctcatggtct	gggttaacta	15060
teestagesag	attacatata	attttaagaa	gcctagaaag	aaaaaattaa	taatgtaaaa	15120
Lygatagaaa	tttagacaca	2+22/22222	ctctcccad	cactgtggct	caaatctgca	15180
ataatattaa	ctigataata	tenactata	ccatcactta	adaccadadt	tcaagactca	15240
atcccagcta	cccaggaggc	cyaggeggaa	ggarcacteg	agaccagagt	ttttaaaaa	15300
gcctaggcaa	cacggcaaga	aactgtetet	aaaaaaalla	aaacttaaat	ccccaaaaaa	15360
gaattctcaa	agcgtcacaa	aaactggaga	ttaaggtaca	ggaagtgtga	aytaatatta	15420
ctatgctaat	ggttttttt	ttttttagaa	aggtataacc	aaaagatttc	ttttttaagt	
cgataaactg	aqaaaqataa	qcatatcttc	caattaacag	agggggagga	aaagccagat	15480
acaacaaaat	aaqatataaa	ttagtttcca	gttgaaaaca	agagtaggag	ttattttgca	15540
tcacctcacc	tqtqacctcc	cccaqcccaa	aaaacactac	tgataaacag	ggtagaaaag	15600
catcatctca	gataaagcag	gaaaaactgc	cacagtctca	aaccacaaac	tataagcaca	15660
cacctggcca	accctgccaa	gtctgggctc	agtaggagga	acgtgctgag	agctaggatg	15720
taccaactta	gacattctgt	gggatacaga	tgtccctgga	agggtcacac	catctcaaag	15780
gcacctgtaa	tqcccactga	ttacagccac	catatgtgag	agagaaactc	agggcactta	15840
gagagtataa	caagaacctt	atgtcatctq	agatgaggaa	tcctcagccc	tgcaaattaa	15900
5-5-5		5	5 5 5 5	-	-	

	agaacaactg	gcaaaacata	aatatccaca	acttttgttt	cagtaattcc	15960
CCaactcttt	atcaatccaa	actacatcac	acaccacata	cacacacaaa	atggtattta	16020
actellagal	gtttataata	ageacacgag	acageagata	catatotete	aataggatac	16080
ctgcagcatt	agggtatgta	gcaaaaaaca	agaaacaacc	aagagtgata	tagatataca	16140
tgggtacatg	taaaaagctg	totattacat	Caaaacaaac	tcaagcagca	acaggataga	16200
cagatggaca	taaaaagctg	charte	72222222	tctactctct	gcacagaaaa	16260
cttatgatag	tcagtatgag	ccaacticcg	cacaatatto	acatotttta	ctaagttgta	16320
catctgaaag	aacagaaaca	thrococcas	cagaacactg	ctttttccaa	aaagaaaatt	16380
tatctatact	gcttgtaatt agcatttctt	tttaccccaa	gtaagaatta	atactastaa	aataaaaaaa	16440
caggaaataa	ageatttett	Laaacttcat	gtttaaacaa	tatacatasa	cacacccttt	16500
ttettattea	tcataaacac	acacagcaca	totttocttt	taccttctat	cccttcactc	16560
acttgataaa	taccatgttg	aatatttag	atatacatta	tacctaaacc	cactttaata	16620
aaaatgcggt	tataaataaa	tgtactttc	acgegeeeee	actoraceta	aaactaccct	16680
taactttaca	gtcccattat	cattatagte	tagagetag	tonantana	acaaccccta	16740
ttcatttgga	accettatta	aaatgecaca	cacagettet	ccaaataaaa	toottataat	16800
ggacctgaca	ctaggcttcc	tttgttgcta	ccataatgg	ccaageeeg	cttattatat	16860
acatcttctt	tcattttatt	gctacatatc	caagggtttt	acacycicc	222ttgagga	16920
cttaattcaa	aacaccatca	egetettte	cagatgaaaa	caayyaaaay	aaattgagca	16980
actgactgac	ttaaaggtca	taaaactata	tagtagcaga	gtcagcaaaa	gaagaaacac	17040
acatctccca	agtagaggct	gaaaaccagt	accattcacc	tccagggtga	getatataca	17100
gattacaaag	tcaccttctc	taaatgttca	aactgaatcc	catacccata	terreseast	17160
acctcgtaag	aacagcctca	gatettatta	tagccttttt	tttagcatgc	Egaagecaat	17220
aaaatgcttc	ccattcagca	agagaaacaa	gttctgaaac	actgaataat	ctgcccaggg	
cctatgaaca	tttccactgt	gagaaatgtt	ctccactgtg	tggagaagat	CCCCACCCC	17280
ctccacacag	gcagaacatt	agaaaaattc	ttggattcta	tgatgcacag	cttaggagtc	17340
tgtttagcac	aatttaagtc	caaatagtta	ttaaatcctc	ctctgttcca	gaaacagtgc	17400
taaatactgt	gaatataaaa	attgaaaaga	tactctcctg	gctcccaaga	aagtcagcca	17460
gatagaggag	acacaggcac	acaaatcact	gtcacatgaa	getetacete	cctaacttca	17520
aacgagggc	taagtcacca	agaatacagt	agcagttgtg	actacgagta	actactataa	17580
ttcaatactt	tatcttccct	taqaaaactc	ttctcccttg	gaaatttatt	tgcatttcta	17640
aataccattc	cttactaaaa	qqaaqcaggg	ctccttgggg	aaatagctga	ttctaggtgt	17700
ggactatgaa	atgaaaatgg	tgagtctggg	acatcccatg	ttgcccagaa	atcaaggaac	17760
tacccaaaga	ttaacagagt	catottaaat	ggacctaaga	gtgaaccaga	aggagctcac	17820
tttgggggg	gtggaacaat	ttcaaqaaaa	acatqacaqt	aatgaattat	aaaacatgaa	17880
ttaaaataca	tattqqtact	aaaaaqaqaa	caaaaggatg	tggctttgga	taaagctctt	17940
cttcatqqaa	gaataccagc	taataaatgt	aaaggaaatg	agagaattag	aaaaattatc	18000
attttqtaaa	ccttaatata	ttcacctaga	catqctaaaa	ccactgagta	aaaggctgct	18060
taggaagagg	atoctcacat	gatctcagag	tttcacacca	cagataattt	attagataca	18120
ggaaggaaga	totoatcaag	cttcctqtqa	cccccagcca	ggccccacaa	cactatgtgc	18180
ctccttgtga	tataaaaact	acacagcatc	qcccacacag	cttctcgcca	aaactgtttg	18240
aagctaatca	caaqqqaaqa	actggacagc	ttctgaccat	gagacgctcc	accagacaac	18300
ttacttaacc	tctccaaaqa	aacttqcttq	qcctctccaa	agaaaactca	gitteatta	18360
aaaacaaaac	taattattta	aaaacaaacq	aaaagcaagt	tgtggacttg	agetecaggg	18420
acagagcaga	catacttttc	cctqttcttc	ccagtaagtg	gtaataaaaa	ccctcaacac	18480
tagatataaa	acaaatataa	gaaggttctg	gaaggggaag	aggaggcaga	ctatccaggt	18540
gccttgaggc	ccacagaaca	acccaqtqat	gggttcactg	ggtcttctt	ttgcttcatt	18600
atctcagact	tagaactgaa	qcaqcaqqca	acttcaaaac	accaaggggc	acagattgaa	18660
aagccccaag	aaaaqcctqc	cctctctagc	caaaggacca	ggaaggagac	agtctaatga	18720
gatggaacac	atttagacag	taactgccca	tttaccagca	ataactgagc	agggagccta	18780
gacttccagt	cttqtqaqqa	cqtaccaagg	tacccaacac	ccccaccaag	gctgagtaag	18840
gactgcgact	tttatccctq	catggcagta	qtaaggagcc	catccctcac	ccgccagcag	18900
tatcagaga	acctggactt	ccactcccac	ccaggagtga	tgaggccctc	cctgctgggg	18960
tcatqtcaqa	qqaqqcctaq	tqqaqattca	qtgacttaac	cttttcccag	agataatgag	19020
gccacctttc	ctccctcttc	ccccatggtg	acagtgaaag	cactgtggca	agcagtaggc	19080
actcctaccc	ctcctaqcca	qqqaqqtatc	agggaggcca	agtagggaac	cagaataccc	19140
acaaccaccc	agcagcaaca	ggggtccccc	accccattgg	gtgtcaatgg	aagcagagcg	19200
gaaagcctgg	atatttaccc	ccatctagaa	qtaacaagct	gatgtccccc	ttcttctact	19260
acaatqqtqt	tcaaaacagg	tttaaataag	gtctagagtc	tgataacgta	atacccaaat	19320
cottgaagtt	ttcattgagg	atcatttata	ccaagagtca	ggaagatccc	aaactgaaag	19380
agagaaaaga	caattgacag	acactagcac	taaqaqaqca	cagatattag	aactacctga	19440
aaggatgtta	aagcacatat	cataagcctc	aacaggctgg	gegeggtgge	tcacgcctgt	19500

				•	•	
aaccccagca	ctttaggagg	ccgaggcagg	tggatcacaa	gatcaggaga	tcgagaccat	19560
cctccctaac	accottgaaac	cccgtctcta	ctaaaaatac	aaaaaaaat	agcaaggcat	19620
catagtagac	acctgtagtc	ccagctactc	gggagcctga	ggcaggagaa	tggcatgaac	19680
9909909990	gaggagtgag	ccgagatcgc	accaccacac	tccagcctgg	gcaacagagc	19740
cegggaagag	gagaagagag	aaaaaaaaa	aaaaaaaagc	ctcaacaaac	aactacaaac	19800
adjacticgt	cocatanaaa	aaaaatcttġ	ccaaacaaat	aaaagatata	tattttggcc	19860
grgcrrgaaa	caaatyaaaa	tgtaatccct	gcadagadac	addagacaca	aggcggatca	19920
aggtgcagtg	getcacagee	rgradicter	gcactteggg	aggergagge	ctctactaaa	19980
cctgaggtca	ggagtttgag	accagcctga	toaatataat	gaaaccccgc	CCCCaccaca	20040
aatacaaaat	tagccagtca	tggtggcaca	tgeetgeaat	cctagecace	caggaggccg	20100
aggcaggaga	atcgcttgaa	ctcaggaggt	ggaggttgcg	grgageegag	acceegecat	
tgcacattgc	actccagcct	gggcaacaag	agcaaaactc	cateteaaaa	aaacagacac	20160
atattttaat	ggaaatttta	gaattgaaaa	atacagtaac	caaattgaat	ggaaagacaa	20220
catagaatgg	agggggcaga	caaaataatc	agtgaacttc	aacagaaaat	aatagaaatt	20280
acccaatatg	aagaacagaa	agaaaataga	ctggccaaaa	aataaagaag	aaaaaagagg	20340
agcagcagga	ggaatgatgg	aaaaagagaa	aggaaggaag	gaagggaagg	agggagggaa	20400
ggagtgaggg	agaaagtctc	aaagacctct	gagactaaaa	taaaagatct	aacacttgtc	20460
atcagggtcc	aggaaagaga	caaagatggc	acagetggaa	acgtattcaa	aaaataatag	20520
ctgaaaactt	cccaaatttq	gcaagagaca	taaacctata	gattcgaaat	gctgaacccc	20580
222122222	cccaataaaa	tccacaccaa	aatacatcat	agtcaaactt	ctgaaaagac	20640
gaacagaaag	aacatcttaa	aagcagtgag	tgaaacaaca	cttcatgtat	aaqqqaaaaa	20700
gaaaagagaa	aacacatttc	ttacagaaat	taaggaagcc	agaaggaaat	gacacaatgg	20760
taattcaagt	atanagarre	agaagtgtca	acacaaaatt	ctagattcag	taaaaatato	20820
tttttaagtg	ccgaaagaaa	tcaagacagt	ctcacatata	ccagacteag	acaatatott	2088.0
cttcaagaat	caatgggaaa	ccaagacagc	angaraaa	atacaacaag	ccasasasta	20940
gccagcagat	ctcccctaaa	ggaatggcaa	aaggaagatt	acgcaacaga	caaaaaaacg	21000
atgaaagaag	gaatccagaa	acatcaagaa	gaaagaaata	acatagtaag	cadadataca	21060
tgtaattaca	ataaaatttc	tatctcctct	Laagacttct	aaaccacacc	gatggttgaa	21120
gcaaaaatta	taaccctgtc	tgaagtgctt	ctactaaatg	tatgcagaga	attataaaty	21120
gggaaagtat	aggtttctat	acctcattga	agtggtaaaa	tgacaacact	gtgaaaagtt	
acatacacac	acacacgtaa	gtatatataa	atatatgtgt	gtatatgtgt	gtgtatatat	21240
atatatacat	ataatgtaat	acagcaacca	ctaacaacac	tatacaaaga	gataataacc	21300
aaaaacaatt	tagataaatt	gaaatggaat	tctaaaaaat	attcaaatac	tctacaggaa	21360
gacaagacaa	aaagagaaaa	aaagaggagg	acaaactaaa	ttttttaaaa	acataaataa	21420
aatggtagac	ttaagcccta	acttatcaat	aattacataa	atgtaaatga	tctaattata	21480
tcaattaaaa	gacagagata	gcagagttaa	tttaaaaaca	tagctataag	aaacctgctt	21540
tagactáaat	gcagtgactc	acacttqtaa	tcccagcact	tcgggaggcc	aaggcgggtg	21600
gatcacctga	ggtcaggagt	tccagaccag	cctggacaac	atggtaatac	cccatctcta	21660
ctaaaaatac	aaaaaaatta	gccaggcatg	gtggcacacg	cctgtagtcc	caactactca	21720
agaagetaca	acacaaqaac	tgcttgaacc	cgggcagcag	aggtagcagt	gggccaagat	21780
tacaccactc	cagectgaae	gacagagtga	gactccacct	cagttgaaaa	acaaaaaaga	21840
aacctocttt	aaatatacca	acatatgttg	gttgaaatta	aaagaataaa	atatatcatg	21900
aaaacattaa	tcaaaagaaa	ggagtggcta	tattaataac	ataaaataga	cttcagagaa	21960
aagaaaattt	саададасад	gaataaaagg	atcaagaaaa	gatcctgaaa	qaaaaqcaqq	22020
casatcastc	attctgcttg	gagattcaac	accetetett	aacaactgat	agaacaacta	22080
dadaccaacc	tracratora	gttgagaaga	acttaacacc	actgaacaac	aggatctaat	22140
gacaaaaaaa	ggaacactct	acccaacaat	agcaaaataa	acattcttt	caagtattca	22200
agacacctac	cottagaccc	taccctgggc	cataaaacaa	ageteactag	tgattgccga	22260
cigaacacac	ccccagaccc	agagetgeat	Cacaaaaaaa	agootaccag	ttaaagagtg	22320
aggerrggar	ggacagugga	agagetgeat	99999999	attetectes	tattaacaca	22380
taggatttct	ttttgggata	atgaaaatgt	tocadaticy	accycygrau	aacatatoot	22440
actctacaaa	tataaaaaag	gccattgaat	tgtacgtttt	ttessagest	adcacacagge	22500
atgtggatta	tatctaacgc	tttttaaaaa	Cttaacacat	tttaaagaat	agaagttata	22560
cagagtgtgc	tctactggaa	tcaaactaga	aagaggtaac	Eggaggataa	cgagaaaagc	-
ctccaaatac	ttgaaaactg	gacagcacat	ttctaaaatc	accegteget	caaagatatt	22620 22680
catttctgat	attcatttt	attgtttaat	gtatttttaa	aaatttctta	agggaaataa	
actgactaaa	aatgaatatg	gctgggtgcg	gtggctcacg	cctgtgatcc	cagcactttg	22740
ggaggccgag	gctggtggat	cacaagatca	ggagttcgag	accagcctgg	ccaagatggt	22800
gaaaccccgt	ctcaactaaa	aaactacaaa	aagtagccaa	gcgcagtggc	gggagcctgt	22860
ggtcccagct	acttgggagg	ctgaggtagg	agaatcgctt	gaacacaggc	agcagaggtt	22920
gcagtgagcc	aagattgtgc	cactgcacgc	cagcctgggc	gacagagact	gcctcaaaaa	22980
aaaaaaaaaa	aaaaagaata	tcaaaatttg	tgggacatag	ttaaagcaat	gctgagaggg	23040
aaatttataa	cactaaatqt	ttacattaga	aaagaqaaaa	agtttcaaat	caatagtctc	23100
	•					_

				2200023300	2200222202	23160
cactcccatc	tcaagaacac	agaayatyaa	gagcaaaaca	aacccaaagc	aagcaaaaga	23220
aagaaaatat	aaaaataaat	cagtaaaatt	gaaaacagaa	acacaacaaa	gaaaaccagc	23280
gaaacaaagt	actgattctt	cgaaagatta	ataaaattga	caaacctcta	gcaaggctaa	
caaacaaaaa	agaaagaaga	cacqqattac	cagttattag	aatgaaagca	taattagaaa	23340
caactctaca	cattataaat	ttgacaatgt	agatgaaatg	gactaattac	tgaaaaaaca	23400
caaattacca	caactcaccc	aatatqaaat	agataattgg	gatagcctga	taactactga	23460
gaaaattgaa	tttotaattt	taacactctt	aaaacaqaaa	cattaaactt	aatattttat	23520
aaatattada	taaggtaatt	ataccettee	ttaacaaata	aaaacqacaa	attattttgc '	23580
aggtalagg	atgtatgtac	totoaaaaat	atcttcagaa	aaatagaact	ttatttgaag	23640
agecaaagag	taaaaaatgt	ttttaactct	саадаадсаа	atatctgggc	ccagatggtt	23700
aataaggatt	attctaccaa	atottaato	aagaagtacc	accaactcta	catagratet	23760
tcactgaaga	accccaccaa	acguitation	aagaattaatt	ttataaaata	catastacte	23820
ttgagaaaac	tgaagagaag	ggaacatete	ceagereare	tacquageg	taggedeacee	23880
tgatactaga	actgtataag	gacagecaet	Cityacacac	tgcctatggg	tatassassa	23940
ctgcaggaac	agtcagaaaa	aaaaaaaaaa	gaagcactgg	acaagggcag	Lacadadaa	24000
gaaaactggg	ccaggtgcag	tggctcacac	ctgtaatctc	agcactttgg	gaggetgaeg	
ctaataaatc	acctgaggtc	aggagtttga	gactagcctg	gccaacatgg	taaaaccctg	24060
tototactaa	aatacaaaaa	ttagccaggc	agggtggtgg	ggaaaataaa	aaggaaaaaa	24120
aaacaaaaat	aaactgcaga	ccaatatcct	tcatgagtat	agacacaaaa	ctccttaaac	24180
teettaacaa	aatattagca	aqtaqaaqca	atatataaaa	ataattatac	accatgatca	24240
agtgggactt	attccagaaa	cacaaatcta	gttcaacatt	tgaaaacaag	gtaacccact	24300
atatgaacgt	actaaagagg	aaaactacat	aatcacatca	atcaatqcaq	aaaaaagcat	24360
ttaccasast	ccaatatcca	ttcatgatac	totaataaga	aaaataagaa	taaaggggaa	24420
attesttess	ttgataaagc	ttacaaaaaa	ctacaaaagc	ttacagetaa	cctatactta	24480
attettegat	actaaatgct	ttacaaaaga	2102000000	aagcaaggat	gttcactctc	24540
atggtgaaaa	actadatycc	coccccacg	ctaacttctc	casascoata	aceaaccca	24600
attgctctta	tttaacatag	ccccgaagee	ctaactcgcg	taatatttac	agadagggda	24660
atgaaagacc	tgcagattgg	caaagaagaa	acadaaccyc	coctgeeege	agatgacatg	24720
attgtctcat	agaaaatgta	aagcaactag	9991499999	gcagragaga	cacgetggte	24780
aaaggatacc	aaatttcagt	taggaggagt	aagttcaaga	tacctattgc	acaacatggt	_
aactatactt	aatatattgt	attcttgaaa	atactaaaag	agtgggtgtt	aagcgttctc	24840
aççacaaaaa	tgataactat	gtgaagtaat	gcatacgtta	attagcacaa	cgtatattac	24900
tccaaaacat	catqttqtac	atgataaata	cacacaattt	tatctgtcag	tttaaaaaca	24960
catgattttg	gccaggcaca	gtggctcata	cctqtaatcc	caqcatttta	ggaggctgag	25020
gcgagcagaa	aacttgaggt	cgggagtttg	agaccagaat	ggtcaacata	gtgaaatccc	25080
gtetecacta	ataatacaaa	aattagcagg	atqtqqtqqc	gtgcacctgt	agacccagcc	25140
acttgggagg	ctgaggcacg	agaattgctt	gaacaaggga	ggcagaggtt	gcagtgagct	25200
agataccact	gcattccagc	ctootoacao	agtgagactc	catctcaaaa	aaaataaaat	25260
agaggataac	ttttcttaaa	tacaaaacaa	ccaagcgcag	taactcatac	ctgtaatccc	25320
accactttaa	gaggccgagg	caddcadatc	acaaggtcag	gagtttgaga	ccaqcctqac	25380
accacticgg	aaaccccatc	tetactasas	aatatataaa	ttagccaggc	atgtgtagtc	25440
caacacggcg	aggaggetga	cccaccaaaa	tarattara	ccaacaacaa	aggttgcagt	25500
teagetaete	aggaggetga	ggcaggagaa	cactegaac	teegtetea	22222222	25560
gttgagccac	cgcactccag	cetgggtgag	agaacyayac	ant to at a	atatttaaa	25620
caaaataacc	taattttaaa	aacactaaaa	ctactaagtg	aattCagtaa	guttutagga	25680
ttcaggatat	atgatgaaca	tacaaaaatc	aattgagctg	gacaaaggag	gattgttta	25740
ggtcagtagt	ttgaggctgt	aatgcacaat	gattgtgcct	gtgaatagct	getgtgetee	
agcctgagca	gcataatgag	accacatctc	tatttaaaaa	aaaaaaaatt	gtatetetat	25800
gtactagcaa	taagcacatg	ggtactaaaa	ttaaaaacat	aataaatact	gtttttaatt	25860
gcctgaaaaa	aatqaaatac	ttacatataa	atctaacaaa	atgtgcagga	cttgtgtgct	25920
gaaaactaca	aaacqctqat	aaaagaaatc	aaagaagact	taaatagcgt	gaaatatacc	25980
atacttataq	gttggaaaac	ttaatatagt	aaagatgcca	attttatcca	aattattaca	26040
caggataaca	ttattactac	caaaatccca	gaaaaatttt	acatagatat	agacaagatc	26100
atacaaaaat	gtatacggaa	atatocaaao	gaactagagt	agctaaaaca	aatttgaaaa	26160
acassastas	agtgggaaga	atcagtctat	ccagtttcaa	gacttacata	gctacagtaa	26220
tcaagactgt	gatattgaca	daddacadc	tatagatcaa	tocaaccaaa	tagagaacta	26280
202220220	acacacaaat	atocccaast	gatttctgac	aaaggtgtta	aaacacttca	26340
ayaaayaagc	atatatatat	acycccadac	tatagagtaa	ttacacatat	atagggaaaa	26400
acgggggaag	atatgtctct	cattadayyy	cycayaycca	aaaataaata	aaccactaaa	26460
agatgaacct	gaacctcaca	ccctacagaa	adallaacic	addatyacte	tanggattaad	26520
cataagatat	acatctataa	aacatttaga	aaaaggccac	geaeggegge	atttace	26580
aatcccagça	ctttgggagg	ccaaggcagg	tggatcacct	aaggccagga	gittgagacc	
agccggatca	acatggagaa	gccccatctc	tactaaaaat	acaaaattag	ctggacgtgg	26640
tggcacatgc	ctgtaatccc	agctacttgg	gaggctgagg	catgagaatc	gcttgaaccc	26700

26760 26820 26880 aggcgaggta aagaattott acacttgatg ccaaactaag atctataagg ccagtcgtgg tggctcatgc ctgtaattcc agcactttgg tcaactagat gaaaggtata tgggaattca 26940 ctgtattatt ctitcaactt tictgtaggt ttgacatttt tttagtaaaa aattggggga 27000 aagacetgae geagtggete acacetgtaa teecageaet ttgggaggee ggggcaggtg 27060 gatcacacgg tcaggagttc gagaccagcc tggccaacat ggtgaaaccc cgtctctacc 27120 aaaaatataa aaaattagoo gggtgtoatg gtgcatgoot gtaatcocag ctactgagga 27180 ggctgaggca ggagaatcac ttgaacctgg gaggtggaag ttgcagtgag ccgagattgt 27240 gccactgcac tccagcettg ggtgacagag cgagactccg tctcaaaaga aaaaaaaaa aaagaatatc aaacgettac tttagaaact atttaaagga gccagaattt aattgtatta 27300 27360 gtatttagag caatttttat getecatgge attgttaaat agageaacca getaacaatt 27420 agtggagttc aacagctgtt aaatttgcta actgtttagg aagagagccc tatcaatatc 27480 27540 actgtcattt gaggctgaca ataagcacac ccaaagctgt acctccttga ggagcaacat aaggggttta accetgttag ggtgttaatg gtttggatat ggtttgtttg gececaecga gteteatgtt gaaatttgtt ecceagtaet ggaggtgggg eettattgga aggtgtetga 27600 27660 gtcatggggg tggcatatcc ctcctgaatg gtttggtgcc attcttgcag gaatgagtga gttcttactc ttagttccca caacaactgg ttattaaaaa cagcctggca ctttcccca 27720 27780 tototogott cotototoac catgigatet cactggittee cottocotti atgcaatgag 27840 tggaagcagc ctgaagccct cgccagaagc agatagtgat gccatgcttc ttgtacagcc tacaaaacca tgagcccaat aaaccttttt tctttataaa ttatccagcc tcaggtattc 27900 27960 ctttatagca agacaaatga accaagacag ggggaaatca acttcattaa aataatctat 28020 28080 gcagtcacta aacaaataag aacaagaggc tccagaagtg ggaagccaat acccagagtt cctacaatac agtatctgaa aagtccagtt tccaaccaaa aaatatatat atacaggccg 28140 28200 gacatggtag cttatgtctg taatcccagc actttgggat gctgaggcgg gcagatcacc ctaggtcagg agttcgagac cagectggcc aatatggcaa aaccccgtct ctactaaaaa 28260 tacaaaaatt agccaggcat ggtggtggat gcctgtaatc ccagctactc gggaggctga ggcagggaat cacttgaacc caggaggcag aggttgcagt gagccgagat cacgccactg 28320 28380 28440 tatatgtgtg tgtgtgtgt tgcgcgcgtg tgtgtatata cacatacaca tatatacata tatacagaca cacatatata tatgaagcat gaaaagaaac aaggaagtat gaaccatact 28500 28560 28620 28680 28740 tccaaataga gaatatcaat aaaggcatag aaattataaa atataataca atggaaattc cggagttgaa aggtagaata actaaaattt aaaattcact agagaaggtt caacactata 28800 28860 tttgaactgg cagaagaaaa atttagtgag acaaatatac ttcaatagac attattcaaa 28920 28980 tgaaaaataa aaagaaaaaa gaatgaagaa aaataaacag aatctcagca aaatgtggca 29040 caccattaat cacattaaca tatgcatact gagagtaccg gaagcagatg agaaagagga agaaaaaata ttcaaatgat ggccagtaac ttcctagatt tttgttttaa agcaataacc tatacaatca agaaactcaa tgaattccaa gtaggataaa tacaaaaaga accacaaaca 29100 29160 gatacaccat ggtaaaaatg ctgtaagtca aaaacagaga aaatattgaa agcagctaga 29220 ggaaaactta taagagaacc tcacttacaa aagaacatca cttataaaag aaccacaata 29280 29340 atagaaacag ttgacctctc atcagaaaca atgaatgata acatatttga agtgctcaaa gaaaaaaaat aaagatteet atataegaea aagetgtett teaaaaatat acateeaaaa ggattgaaac cagggtettg aagagttatt tgtacateea tgtteatage ageattatte 29400 29460 acaatagcca aaaggtagaa gcaacccaag ggtccatcga caaataaata aaatgtggta tatgtataca caatggaatt tattcagtat taaaaaggaa tgaaattctg acacatgcta 29520 29580 29640 caacatggct aaaccttgag aacactatgc taagtgaaat aagccagcca caaaaggaca 29700 aataccatat tacttcactt gtatgaaata cctagggtag tcaaattcag agatagaaag taaaacagtg gttgccaagg gctgagggag ggagtaacgt ggagttattg ttgaatgggt 29760 29820 acagaatttc agttttgcaa gataaaaaga gttctggaga cagatggtgg tgagggtggt 29880 acaacaatac aaatatactt tatactactg aacagtatac ttaaaaatga ttaacatggt 29940 gaaaccccgt ctctactaaa aatacaaaaa aattagctgg gtgtggtggc gggcacctgt aatcccagct acttgggagg ctgaggcagc agaattgctt gaaaccagaa ggcggaggtt gcagtgagct gagattgcgc caccgcactc tagcctgggc aataagagca aaactccgtc 30000 30060 30120 tcaaaaaata aaaaataaaa aaaatttaaa aatgattaag caggaggcca ggcacggtgg 30180 ctcacaccta taatgccage actttgggag gccgaggcag gcgatcactt gagaccagga gtttgagacc agcctggcca acatggcaaa accctgtctc tgctaaaaat acaaaaatta 30240 30300 gccaggcatg gtggcatata cttataatcc cagctactgg tgagactgag acacgagaat

tacttgaacc	caggaggcag	agattgcagt	gagtcgagat	cgcgccactg	aattccagcc	30360
tagacaacaa	agcaagattc	tatataaaa	aaacaaaaac	aaaaacaaaa	agcaaaacca	30420
aaaaataatt	aagcaggaaa	cgagattgct	actgaggagg	aqaaaqatqt	gcaggaccaa	30480
gagagagaga	agcacaaaac	ttttcaaaaa	atotttaato	attaaaatgg	taaattttat	30540
atatatata	ccacaaaaaa	aagggctggg	gggcaggaaa	tgaaggtgaa	ataaagacat	30600
acgeacceta	caaaagtaga	gaatttgttg	ccttagaaga	aacaccacag	gaagttcttc	30660
cccagagaaa	caagtgaccc	cacacccataa	tctcaattct	cacagaaaat	tgaagcatag	30720
aggergaaaa	tattctgtaa	ctatgaggeau	aacaatocat	attttttcct	ttetteteta	30780
cagtaaaggt	aaaagcaatt	ccatgacatt	ttatatataa	agectattgt	tgaacctata	30840
adatgattta	aaatatactt	gtacaaaata	tocasatasc	tocacaaaaa	agagttggaa	30900
acatatatay	actaggctaa	agaaattact	acadatadta	aagtaatata	acagggaact	30960
caaagetgtt	attttaaaaa	agaaaccacc	acagacagca	acaataatat	aattaaattt	31020
taaaaataaa	tagacataat	accidadaac	cacaactaca	deddeddede	atagaactac	31080
gtaatattaa	cattttggta	tetaaetae	attaaattat	aaatatgaag	tatattctoo	31140
ataggaataa	cacacatgtt	cccaactaga	tactaaaaaa	taactcacat	aaatacagta	31200
taagttaaga	taaaataatt	aaacccctaga	tattagtttg	ctcaccacac	acteacasac	31260
aaaaaataaa	taaaacaacc	addatgtttg	actagette	teteccaatt	ctggaggcta	31320
taccacaaat	tgagtggctt	aacacaaccc	adacycactc	taaaaactat	accasacaat	31380
aacacctgca	atcaaggtga	gtacagggcc	teteretare	cotageteet	ccttaactta	31440
cctcccttgt	ctcttccagc	ttccagtggt	cccagraac	anagegee	ccccggcccg	31500
tagctatatc	attcctagca	accagaaaga	agaaaacaac	adagattaty	gtaaaaaaaa	31560
atgaaatcaa	aaggagaaaa	atggaaaaaa	ataaataaa	ccaaaagcta	accetted a	31620
aagatcaacc	aagttaacaa	accttttaac	tagactgaca	aaaaggaggt	aagactcaaa	31680
ttactagaat	cagaaataaa	agaggggaca	ttactaatga	gggattagaa	aagaacacca	31740
cgaacaaatg	tgtgccaaca	aattagaaaa	cttagatgaa	atggacaggt	teetaggaca	
acatcaacta	ccaaaattta	ctcaagaaga	aagagacaat	ttgaatgagc	tataacaagg	31800
gaagagactg	aattgacaac	caagaaacta	tccacaaaga	aaatcccagg	cccagaagat	31860
ttcactgtga	aattctttca	aacttataaa	tataaattaa	catcagttct	tcacaaactc	31920
ctccaaaaaa	aagaacagat	ctctatttac	aggcgatacg	atctttagaa	aatcctaagg	31980
gaactactaa	qacactatqa	taactgataa	acaagttcag	caaggctgca	ggacagaaaa	32040
ccaatataca	aaaatctatt	atatttctat	acacttgcag	tgaacaaccc	aaaaatgaga	32100
ttaagaaaat	aattcaattt	acaataacat	caaaaagaat	aaaaacactc	aaaaacaaac	32160
ttattcaagt	aagtgcaaaa	cttatactct	agaagctaca	aaacactgtt	aaaagaaatt	32220
aaaggtttac	ataaatgaaa	aactatccca	tgttcatgga	tcaaaagact	tattactggc	32280
aatgetetee	aaattgatct	ataaattcaa	caaaatcctt	atcaaaatcc	cagatgaggc	32340
tgggggtggc	ggttcatgcc	tgtaatccca	gcactttggg	aggctgaggc	acgcagatta	32400
cctgaggtcg	ggagetegag	atcageetga	ccaacatqqa	qaaaccctat	CCCCCCCaaa	32460
aatacaaaat	tagtcagggg	tootoocaca	tacctataat	cccaqctact	cqqqaagcig	32520
aggcaggaga	atcocttoaa	cccaggaggc	agaggttgca	qtgagccaag	accgtgccat.	32580
tacactccaa	cctgggcaac	aagagcaaaa	ttccatetea	aaaaaaaaa	adadadatt	32640
ccagatgact	tcactgttga	aattgaaaag	attattctaa	aattcacatg	gaattgcaag	32700
accttgagaa	tagccaaaac	aaacttqaaa	aacacgaaca	aaatatagga	tgactcactt	32760
gccaattgca	aatgttacga	cacagcaaca	gtaatcaaga	ctgtgtggta	ctggcaaaag	32820
acacatacat	acatacatat	caatqqaata	taattgagag	tacagaaaca	agectaaaca	32880
tctatggtaa	gtgcttttct	attttttct	tttttttt	cttttttgta	gagatagaat	32940
ctcaccatgt	tgcccaggct	ggtcttcaac	ttctgggctc	aagcaatcct	cccactgtgg	33000
cctcccaaag	tactaggata	actoocatoa	gccaccacat	ccagcçcaga	tgattttcaa	33060
aaaagtcaac	aagaccattc	ttttcaacaa	ataggtctgg	gatgatcaga	tagtcacatg	33120
aaaaaaaaaa	tgaagttgga	ccctccatca	cactaaagtg	ctgcgattat	aggcatcagc	33180
caccacatcc	agcccaaatg	attttcaaaa	aggtcaacaa	gaccattctt	ttcaacaaat	33240
aggtetggga	taatcagata	gtcacatgaa	aaaaaaaatg	aagttggacc	ctccatcaca	33300
ccatatocaa	aaattaattc	aaaaatgaat	tgatgactta	aacgtaagag	ttacgactgt	33360
aaaactctta	gaaggaaaca	tacqqqtaaa	tcttaaaqac	gttaggtttg	acaaagaatt	33420
cttagacatg	acaccaaaaq	catqaccaac	taaggtaaaa	tagggtaaat	tgtacctacc	33480
aaaatgaaaa	acctttqtqc	tqqaaaqqac	accatcaaga	aatggaaagc	caaaatagcc	33540
aaggcaatat	taagcaaaaa	gaacaaagct	qqaqqcatca	tactacctga	cttcaaagca	33600
acagtaacca	aaacaqcatq	gtactagtag	aaaaacagac	acatagacca	atggaacaga	33660
ataaaqaacc	caaaaataaa	tccacatatt	tatagtcaac	tgatttttga	caatgacacc	33720
ccttcaataa	atgatactag	gaaaactgga	. tatcgatatg	cagaagaata	aaactagacc	33780
cctatctctc	accatataga	aaaatcaact	cagactgaat	taaagacttg	aatgtaagac	33840
ccaaaactat	aaaactactg	gtagaaaaca	taaggaaaaa	cgcttcagga	cattggtcca	33900
	_	•				

occasagato	ttatogctaa	aacctcaaaa	acacaggcaa	caaaaacaaa	aatggaaaaa	33960
tagcacttta	ttaaactaaa	aagctcctgc	acagcaaagg	aaacaacaga	atgaaaagac	34020
aacctotaga	atgggagaaa	atatttgcaa	actatccatc	catcaaggga	Clagialica	34080
gaccegeaga	grgactaaaa	caactcaaca	gcaaaaaagc	aaataatctg	gtttttatat	34140
gaacacacaa	tctgaataaa	cattctcaaa	ggaagacata	caaatgtcac	tatcattctg	34200
gggcaaaaga	actototoa	ttacttotta	gtgtataaat	ttttaaattq	ggaagtgtga	34260
ctagtactac	actitionic	totttttcaa	gtttgttttg	actattctag	gageettgea	34320
gccaccccac	2002320330	tatasasasa	tgctcaccat	cactaatcat	cagagaaata	34380
agraraaaar	ccactatgag	atatectete	actccagtta	gaatggctac	tatcaaaaag	34440
aaaaccaaya	atacatacta	gcasagattt	ggagaaaggg	gaactcctat	acactgtggg	34500
acaaaatata	acggacgccg	gcaaagaccc	gaaaataata	ctgaggtttt	tcaaaaaact	34560
tagggatgca	aactygtaat	ggccaccacg	cctactactg	ggtatttatc	caaaggaaag	34620
gaaaatagaa	ctaccatatg	tatatagaac	ctcatgttaa	ttgcaacact	gttcacaaca	34680
aagtcagtat	actgaagaaa	tacatgcacc	tcaacagatg	aatggataaa	gaaaatgtgg	34740
gccaagacag	ggaataaatc	caaatgtgta	attacagacg	aatgaaatcc	totcatecca.	34800
catatacact	caacagaaca	ctatteagee	attaaagaag	taagtaaacc	acaaaaaaat	34860
gcaacatgga	tgaacctgga	ggacattata	tttaatgaaa	aaataaagt	cacacaatta	34920
aaacagtaca	tgttctcact	cagacatggg	tgctaaaaag	addatatataa	cacagaacca	34980
gaagggagg	cttgggaaaa	gttaatggat	aaaaatttac	agctacgtaa	tatacatatt	35040
ttttagtgtt	ctatagaact	gtagggcgag	tatagttacc	aacaacccac	tattatat	35100
caaaaagcta	gaagagattt	tggatgttcc	cagcacaaag	gaatgataaa	atcasattat	35160
gatggatatc	ctaattaccc	tgattcaatc	attacacatt	geatacatgt	accaaaccac	35220
cactctgtac	ctcataaata	tgtataatta	ttacgtcaac	aaaaaaagga	adadadayaa	35280
aattaagaca	acccacataa	tggaagaaat	aaaatatctg	caaattatat	acaccigaca	35340
aatatttaat	atttataata	tataaagaac	tcctacaact	caagaacaac	aacaaaacaa	35400
cccaattcaa	aaatgggtaa	aagccttgaa	tatacactta	tetaaagaet	atatacaatt	35460
ggccaataaa	gacacgaaaa	gatgctcaac	atcactagtc	atcagggaaa	Latadattaa	35520
aaccacaato	tagaatgtag	acaccacttc	atatqcacta	qqatqqctaq	aacaaaaayy	35580
taataacaaa	tattaataaa	. gatgtgaaaa	aatcaqaaac	ctcattcgct	gctgttggga	35640
atataaaata	atgcagccac	tttqqaaaac	agtctggcag	CECCECadae	Lactadatac	35700
agagttaccg	tatgacccag	gaatattcct	cctqqqtcta	taaccaaaaa	aacgaaaaca	7 -
tatatocaca	taaaaacttg	tacatgggca	tttataqcaa	cattattcat	aacagcaaag	35760
gtggtaagaa	cccatatgcc	catcatctga	tgaacaggta	aataacatgc	ggtattatee	35820
atacactaga	atattatctq	cccatacaaq	gagtgacatc	cagctacatg	ctacaaggat	35880
gaatict.cgga	aaccttatgc	taaqtqaaaq	aaqccagtca	caaatgacca	cagallalga	35940
ttccatgcat	cogaaatgac	cagaataggg	aaatctataq	agacagaaag	tagattagtg	36000
attagataga	getaggagga	caggtagtac	actactttcc	cagaactact	ggaacaaagt	36060
accacaaact	ggggagetta	aacatagaaa	ttgatttcct	cacagttctg	gagactayga	36120
ctctgagatc	aaggtgtcag	cagagetggt	tetttetgag	qqccctgagg	caaggetetg	36180
teccagget	ctctccttgg	ctaacaaata	gccatcttct	ccctgcgtct	Ecacattate	36240
- ttttctctat	grafacccat	otccaaattt	tgattggctc	attetgggte	alggecaatt	36300
octatocaca	aagtgaagtc	tacttccaaa	aqaaqqqaaq	agggaacact	gactaggcta	36360
aacttatagt	cattttaatg	teegetttte	ctatgagatt	gtgaacacac	agaagtayyy	36420
rrtttatcta	cattotocaa	agtttaataa	gaaaaataga	attcaagaga	agcagttcaa	36480
taggaggaat	ttaatatooo	aactaattac	aaqqtttaqq	gcaggactaa	aaagccagcc	36540
gagataataa	gccaacccag	agattagcaa	cagtgggacc	ccatctacct	accacccatg	36600
aagctggaag	gataaaggag	gggctattat	caqaqtccac	aagccagtgt	cagageceee	36660
gactagaact	gggaccaccc	tagagacact	gtgcaaagca	gaaaacaagg	gggaaaaacc	36720
ctgacttctc	cettestess	acctttcaat	ctcccactag	tgcttcctac	tagccatact	36780
taaccaaaaa	cagtgacaag	gaacactgca	aaatqaaqtt	tgtaggaatc	ateteette	36840
gagacagaga	aatatqqaaq	ggtagaaaat	gaatcagagg	ataaagagaa	aaaaccccga	36900
gractatett	atttatcttt	gtatctccag	tqcctaatct	gtctctcaaa	aaayyaaayc	36960
aattgagaga	aactgaaaac	tccaattgaa	atqaaaqaat	ggagaattac	tggactagaa	37020
G2G22G2G22	aaatttatto	cacatagaat	aaacaagaat	ggattcacaa	aggacgigal	37080
caatcaaaag	ctataatcad	caaagattto	ccagagaaat	taaaaagtgg	taaactcagc	37140
cacqctqtac	aacctgaagg	cacaatqcat	gaaaacqttt	caagaaacga	. Caayattiya	37200
actrazattr	taagtgettt	tecagaatet	ctcaaqacqa	ttatatagct	accidation	37260
artaaataaa	atggaaactt	actaaacttt	ccccttqtat	taaactaaca	tatgtcclaa	37320
tagcaaacga	. ttctggaatt	cctagagtaa	aatatattt	gtcaaagtgt	attgetett	37380
taatattctd	ctoacctcct	tttqctattt	aggatatttq	tatacacacc	acacytaaat	37440
ttggtctata	gtttacatct	acqqqcttat	actqttcttt	ttttcattt	tttaaaattt	37500

```
ccaaccccca gtatccatat actgctctct atcagggtta ttttaacttt gtaaaatcag ctgagatgct ttccatgttt tttttttta ttttctgcca catttgaata gcataggagt taccaccatc aaccttggat tatttaagca ttcacgattc cacgtgtgga ttttttattc
                                                                                    37560
                                                                                    37620
                                                                                    37680
                                                                                    37740
agagtettte ttgtcattce tgctatcage acagaaccca atetcagett tecagetata
                                                                                    37800
ctctcacccc atggaatttg cagatgaagt tcaaaaggac ctttgcatta tcctgcctcg
coetettece cetteattta gacateacet tettetagaa egtettacet gacatgeeet
                                                                                    37860
gctcccaacc cctgctgccc aattgtgtgc tctcccgtgt cctggcctgc catcctcttt agtaattgcc tgctccctca tctgtctccc cacccagaca ttaagctgaa tagactggat
                                                                                    37920
                                                                                    37980
tigtgtcitg tccatcacta taaictcagc acctagiacc tagtaggiac ttaccaigta
                                                                                    38040
ttcattagca aaatgttatg tataaccttg caccttaaaa acaagagaag gaagacaaaa ttaagtctta agactatggt ttagaacatg gatcagaaac tacagtctgc agcccaaatc
                                                                                    38100
                                                                                    38160
cagaccaaat gaagagacca tgttcattta catacaacct atagcagctt tcacactaca
                                                                                    38220
ggagcagagc taagtagttc caagggaaca cacggccctg caaagcctaa aatatttact
                                                                                    38280
                                                                                    38340
ctatagetet teacagaaaa agtttteaga teeetegttt agaactettg tteatatgea
                                                                                    38400
atttcactaa accatagttt titgggttig titggtittt titggcaaaa aggaatgagc
cgatccagaa aaggttgaaa agaatgaatc attactgctg aaagaatgtg cacacagtcc
                                                                                    38460
gtcagtattc tgctgccatg ctgacaccca tccaatagtg tcatgagatg cagcagctac tactgtgttc tcaatgccga gtccacccac tccataacca tgtccaagca atcttgggaa
                                                                                    38520
                                                                                    38580
                                                                                    38640
catcatcacc atgettgttt atcettaagg tattgeetca catacagcag tggetggtca
taaagtcaaa tgacactagt ggccaggagg tcaagagaat gagtgaggac aggtgggtag gcagcccagg ccctagcaac agcaggagct caccctcag tcactctagc caggactgaa
                                                                                    38700
                                                                                    38760
atacttttca ccctttcaag agagactagg aatctggatt tttatgtgaa atatcttgat
                                                                                    38820
                                                                                    38880
tactaaatgt tgtcaacaga catgtcaaaa ggtaaaacta agtaagttca tggggcagat
tgactattca ggttatagaa ttaaggattc ttatccaaca cagataccaa ccaaaaagct
                                                                                    38940
gacgtataac atattaggag aaactatgtg cactgtegaa acatcaacaa ggggetaatg
tetaaaatag tetatattgg attecagttg aaacatgggg aaaggacatg aacaggcaac
                                                                                    39000
                                                                                    39060
                                                                                    39120
ttatgtcaat ggaaactcaa aaagataaca agcatatata aaagcattct caaattcagt
agtaaacaga cagatgcaaa taaaaagagg gaaactgctg ccgggcacag tggctcacac
                                                                                    39180
                                                                                    39240
ctgtaatccc agcactttgg gaggccgagg cgggcggatc atgaagtcag gagatcgaga
ccatcctggc taacatggtg aaaccccgtc tctactgaaa acacaaaaaa ttagccaggc
                                                                                    39300
gtagtggtgg gcaccagtag tcccagctac tcaggaggtt gaggcaggag aatggcatga
                                                                                    39360
acccaggagg cggagattgc agtgagccga gaccatgcca ctgcactcca gcctgggcga
                                                                                    39420
                                                                                    39480
gtaaataaat aaaaagagag agactgctaa agtctagaaa gttgaatgat gccaagcgca tgcaaagatc agggccttgg gatggccggg tgcagtggct cacgcctgta atcccaccac
                                                                                    39540
                                                                                    39600
tttgggaggc caaggegggc ggatcatgag gtcaagagat caagaccatc ctggccgaca
                                                                                    39660
cagtgaaacc cggtctctac taaaagtaca aaaaaatata tatatata tatattatta
                                                                                    39720
tattatatat atatatatca gagccttggg aatccttgtg tgctgctggg gaaggtagtg gtgcagccac ccttgacagc aatctggcag tacttggtta tattaagtat aggcacacac
                                                                                    39780
                                                                                    39840
cacgaccagg cagtectact cetgggteta aateccaaag aattetcaca caagtecata
                                                                                    39900
aggagacatg tacgaggete atteageatt actgggagtg ggaateaace tgggtgteea
                                                                                    39960
                                                                                    40020
tctacaggag acgagatgga caaaatgtgg tggatattaa gaccagaatc accaagtaac
agagatgggt ggtgagtgac aatcctaaga tacagaataa aggctagaac atgatgccat tcatgtaaat taaaaataga tgcacacaaa gcagtatacg cgtgaccctt gaatagcaca
                                                                                    40080
                                                                                    40140
ggtttgaact gectgtgtee acttacatgt ggattttett ceaettetge tacceccaag acageaagac caaecectet tetteeteet eeceeteage etacteaaca tgaagatgae
                                                                                    40200
                                                                                    40260
                                                                                    40320
aaggatgaag acttttatga taatccaatt ccaaggaact aatgaaaagt atattttctc
                                                                                    40380
ttoottatga ttttotttat ototagotta cattattota agaatatggt acataataca
catcacacgo aaaataaatg ttaattgact gtttatatta tgggtaaggo ttocactcaa
                                                                                    40440
 cagtaggetg teagtagtta agttttggga gteaaaagtt atacacagat tttcaactgt
                                                                                    40500
                                                                                    40560
graggraatr agttrocctg according tgttracggg traactgtat atacaraaaa
                                                                                     40620
gtattatatg aacctcatta gaatagctgt ctatagggag aagagaatga gagtgggata
                                                                                    40680
 aaacggaatg aacaaataaa ccaacaaatg cattaacaag caaaacaaca gaggggcttg
catgggccag tgatgataaa gggctaagaa tgagaatata attaattcaa ttcctcacac
                                                                                     40740
 ctgaggteta aaaccaagga aagggaggge caggegtgga ggeteaegee tgtaateeca
                                                                                     40800
 gcactttggg aggctgaggc gggcggatca caagattagg agtttgagat cagcctggcc
                                                                                     40860
                                                                                     40920
 aacacagtga aagcccatct ctacaaaaaa tacaagaatt acccaggtgt ggtggcacat
gcctgtägtt agctactctg gaggctgagg caggagaatc acttgaaccc aggaggcgga ggttgcaggg agccgagatc acaccattgc actccagcct gggtgacaga gtaagactct
                                                                                     40980
                                                                                     41040
                                                                                     41100
 gtotcaaaaa aataaaaaaa ataaaaaaac agagaaaggg aggaaactag atccaggotg
```

actagataca	gcctttagag	ttagaaaaga	tgatttgaca	atctaagccc	acactcagat	41160
tgaatgaaat	tgaaaagcct	ttcaaactaa	aacatttaat	tacaccatct	gctgcagaca	41220
gaactcagac	aactcaaaca	ggtaatgtca	acataatatt	ttatatcacc	acceteaaca	41280
cagastagas	atcagctgca	tatasacsa	toactagaat	gaagaaagg	ctgcttctta	41340
cagaacaaaa	gtggttcttt	ccassacst	taataggcac	cagetetate	catgtcaccc	41400
tececcia	catggggtat	ataactat	cttactatta	attectease	gaattcccaa	41460
tgcagggaga	gattatacac	acaaccacga	casasactat	ccastasasc	cacccccaa	41520
tettgtggaa	gattatacac	aatgaggcaa	caaaaactat	ccaacaaaac	taccesttta	41580
aagccagtga	caaagaagcc	agigalgaaa	tasaaataaa	taagageega	agetageage	41640
gggaagaaag	accaacatgg	acgggggtga	-cagggtggt	tetetetete	agctggaaga	41700
gaagtggcag	atctctgagc	tggatgatgg	gecaetaeca	cotgcatatg	gctaattaaa	41760
gaccatgtgt	ggattttta	ttcagctctt	tegtgteatt	cetgetatea	geacagaace	41760
caatctcaac	tttccagcta	tattgagcta	aactteteae	cccacggaac	ttgcagataa	
agttcaaaag	gatectigec	ttttcaaaat	aattttgaat	ggccgagcag	recettetgtg	41880
ctctctcact	gacaccctct	caaggctgct	gagcacgcgc	catgetatgg	Ctttctccaa	41940
catcaggaaa	tgttctccac	tcagtttcac	cttaatacaa	atgtgttctc	tcttcagaga	42000
aggcaaaaaa	attcatgacc	atctgactgg	gagaagtcat	ttctaggtaa	agtgtccatc	42060
tttttctgag	gaacacagga	ggaaaatctt	acagaaaaga	gttaacacag	caggcctaag	42120
actgcttttt	aaaataaata	aataaataaa	taaataaata	aataaataaa	taaataaata	42180
aataaatgaa	tgatagggtc	ttctqtattq	gccaggctag	tctcaaattc	ctggcttcaa	42240
gagatectec	caccttqqtc	tcccacagtg	ttqqqattat	agacatgagc	cattgtgctt	42300
ggcccaagac	tgttattctt	aaaaagtctc	ataaaaagca	tggttaatcc	ttggctggca	42360
cctgggaact	tagatttcag	aagggttccc	accatccaac	ctggaaagag	ggactcactg	42420
tocctaaatt	attgtgtggt	ttatqctqaa	ctcctqcttt	tcttcaggta	gcgtggaatg	42480
tagtatatac	tgggcaaagg	gagectacat	gaccagcccc	caataaaaac	cctgggtgtt	42540
gggtctctag	tgagtttccc	togtagacag	catttcacat	gcgttgtcac	agctccttcc	42600
tragggagtt	aagcacatac	atcctgtgtg	actocactoo	gagaggatgc	ttggaagett	42660
atacctaact	tcctttggac	ttggccccat	gcacctttcc	ctttgctgat	tatacttat	42720
atcetttcae	tgtaataaat	tacagccgtg	agtacaccac	atactgagtc	ttccaagtga	42780
acceccacat	ctgagcatgg	tectgagaa	ccccaacaca	gaaataaatt	ataaaagacc	42840
accaccagac	catggtggcc	cataccagta	atctcagcgc	trtgggaggc	cgaggcagga	42900
aaggactggg	agcccaaaag	ttcaaactta	cactoaccta	tgactgcgcc	aatgcactct	42960
ggaccagcca	acagagcaag	cccaaagcca	cagegaceta	2201220300	atacttctcc	43020
aacctgggag	tcttaaaatt	accordicate	ctacaacaca	ttttaaaaca	ctasatcasa	43080
etteçaagig	CCCCaaaacc	taatggaatg	gcagaacac	ctctaccaaa	tttctgaaaa	43140
agaaacctgg	aaaacaagag	cgccgacggc	caaccaaaac	gcccaggaaa	agettaggag	43200
gtaaaaagta	ctcagaacca	gattacctga	gcaaaccata	geccaacaca	taacaccaca	43260
gaggctgtta	tgcagaagga	aatggtaaca	ggtttccagg	aacagacttg	tetesstess	43320
tagaacagca	gaggtagaac	ctgacaaggt	gattacctgg	ggaactgcag	cccgaacgac	43380
caggactgtt	ggaccettee	cctcacatgg	aatacacacg	ccacccagca	gcacaccaca	43440
getetteaac	aatcacagga	ggcacgctac	gectagtaag	acaggaaaaa	aggaattete	43500
aaacttcgaa	gatgaacaca	taaagaatca	ccaagtttt	attcagtatg	atgaaacagg	
gacactgaat	caacagaaca	caaacccaag	caaagataat	tactagagca	catagaagaa	43560
attattagat	attctťggga	agacctaagg	ggacattata	aagagcaagc	agttggtatg	43620
tgacgatctt	tgtgatatac	.caagaaataa	aaacacagga	tgaagaccag	atagagaata	43680
atgctactat	ttqtqcaaaa	aaqqaqaaat	ggagaatctg	attcatattt	gcttgtattt	43740
gcatgaagaa	actttggaag	gtacataagt	aactaacaac	aatggttacc	tacttgtaag	43800
gcgagagaag	taagaggaca	ggaatggtgg	gaacaccttt	tgtgtccgga	attggtgggt	43860
tcttggtctg	acttggagaa	tgaagccgtg	gaccctcgcg	gtgagcgtaa	cagttcttaa	43920
aggcggtgtg	tctqqaqttt	gttccttctg	atgtttggat	gtgttcggag	tttcttcctt	43980
ctaataaatt	cataatctca	ctgactcagg	agtgaagctg	cagaccttcg	cggcgagtgt	44040
tacagetett	aaggggggg	atctagagtt	gttcgttcct	cctggtgagt	tcgtggtctc	44100
gctagcttca	qqaqtqaaqc	tgcagacctt	cqaqgtqtqt	gttgcagctc	atatagacag	44160
tqcaqaccca	aagagtgagc	agtaataaga	acgcattcca	aacatcaaaa	ggacaaacct	44220
tcagcagcgc	ggaatgcgac	egeageacgt	taccactctt	ggctcgggca	gcctgctttt	44280
attetettat	ctggccacac	ccatatcctq	ctgattggtc	cattttacag	agagccgact	44340
gctccatttt	acagagaacc	gattggtcca	tttttcagag	agctgattqq	tccattttga	44400
cagagtocto	attggtgcgt	ttacaatccc	tgagctagac	acagggtgct	gactggtgta	44460
tttacaatcc	cttagctaga	cataaaggtt	ctcaagtccc	caccagactc	aggageceaq	44520
ctggcttcac	ccagtggatc	coccatcaot	gccacaggtg	gagetgeetg	ccagtcccqc	44580
accetacace	cgcactcctc	agccctctgg	tagtcgatag	gactaggege	cgtggagcag	44640
gagatagtac	tgtcagggag	acticagacca	cacaggages	caggaggtag	gggtggctca	44700
שבשיםכבב	-233343	20-0333003				

WO 01/27857

_						44760
ggcatggcgg	gccgcaggtc	atgagegetg	ccccgcaggg	aggcagctaa	ggcccagcga	44760
gaaatcgggc	acagcagctg	ctggcccagg	tgctaagccc	ctcactgcct	ggggccgttg	44820
gaaccaacta	accaaccact	cccaatacaa	ggcccgccaa	gcccacgccc	accoggaact	44880
9990099009	5005500500	~~~	seasttees	accacacata	taastaasa	44940
cacgetggee	egcaagcacc	gegracagee	ceggeeeeg	cccgcgcctc	CCCCCCACA	
cctccctgca	aagctgaggg	agctggctcc	agecttggee	agcccagaaa	ggggccccca	45000
cagtgcagcg	gtgggctgaa	gggctcctca	agcgcggcca	gagtgggcac	taaggctgag	450,60
#3####################################	dadedadeda	ggactgccag	cacgetgtea	cctctcactt	tcatttatgc	45120
gaggeacega	3030303030	554665665	attatatta	atatttt		45180
ctttttaata	cagtetggtt	Ligaacacig	accaccicac	ctatttttt	LUCULULULU	
tgagatggag	tegetetetg	tegeceagae	tggagtgcag	tggtgccatc	ctggctcact	45240
gcaageteeg	cctcccaaat	tcacaccatt	ctcctgcctc	aacctcctga	qtaqctqqqa	45300
				tattttttt		45360
						45420
				acctcgtgat		
cggcctccca	aagtgctggg	attacagacg	tgagccactg	cgccctgcct	atcttaccta	45480
tttcaaaaqt	taaactttaa	gaagtagaaa	cccgtggcca	ggcgtggtgg	ctcacgcctg	45540
				aggtcaggag		45600
taaccccagc	200000000000000000000000000000000000000	9009499095	505500000	-550005545	accacacata	45660
teetgyttaa	Cacagogaaa	coccyccycc	accaaaaaca	caaaaaatta	geegggegeg	
gtggtgggca	ccggcagtcc	tegetactgg	ggaggctgag	gcaggagaat	ggcgtgaacc	45720
tgggaggcag	agcttgcagt	gagccgagat	agtgccattg	ccttccagcc	tgggcgacag	45780
agroagacto	cacctcaaaa	aaaaaaaaa	aaaatagaga	cccggaaagt	taaaaatatg	45840
2505252000	*********	atanagaan	aggetaaag	attascaass	casatctass	45900
ataattaata	LLLadadaca	Cicaagagai	gggccaaaga	gttgacggaa	Caaacccaaa	
tattagattg	gtgacctgca	aaaccagccc	aaggaacatc	ccagaatgca	gcccataaag	45960
ataaaqaqaq	catttccgct	gggcacagtg	gtatggcagg	ggaattgcct	gagtccaaga	46020
				tgggcaacac		46080
						46140
				tatttgagag		
ttatttcaag	aaacatgaaa	gataaatcaa	gatattctaa	ttcccaagta	agaataattc	46200
cagaagcaga	aaatagaata	gaggcaagga	aacactcaaa	acttctccag	tgccatagaa	46260
				tgagcaggaa		46320
astassatat	taagggagtg	taatacccaa	acacaataac	tcatacctat	aatcccacca	46380
gatteactet	caagecageg	tggtgtttaa	gegeagege	tcatgcctgt	aaccccagca	
ctttgggagg	ccgaggcagg	tggatcacct	gaggtcagga	gtttgagatc	agreaggeea	46440
acatggtgaa	accctgtctg	tactaaaaat	acaaacatta	gctgggtatg	gtggtgcaca	46500
tctgtaatcc	caactacttq	ggaggctaag	gcaggagaat	cacttgaaac	caggaggtgg	46560
aggttgtagt	gageegagat	catoccacac	teceageeta	ggtgacagag	caagattcca	46620
						46680
				attttagaac		
aagaaaaaaa	attgtaaagc	ttcagagaaa	ataaacatta	actacaaaga	aacgagagtc	46740
agacgcgtgc	acttcttcct	agataccage	agataaagca	atatctccaa	aattcagaag	46800
offthaacot	agaateetat	acceagtesa	gaatattcac	atggaaaagt	gaaataaaaa	46860
200000000	25222222		atttaggatt	casasatas	ctcaaaacaa	46920
acattgttta	aacacgcaag	ggcccagaaa	gettactact	cacagaatcc	CLyaaaacaa	
aaccaaataa	tcacttaagg	actcattaag	aaaacaaatg	aaataaaagc	accaatgatg	46980
aqtaaataat	cagaaaaatt	tacagtttac	ctaaataact	gtttatgcat	aatgtatgaa	47040
aacccaaaaa	tttaatatgg	gacagaatta	aaatcatgat	aagattcttt	tttgctttac	47100
						47160
				agagaaaaaa		47220
atgtgtgaaa	aactaagggt	accaaaacag	tgcaaattca	tttatcatca	ggaaaateea	+
aattaaaacc	acagtatcca	ccagaataac	taaaaggtaa	aagacagaaa	ttaccaagag	47280
ttqqcaaqaa	tqtqqaqcaa	ccacatatac	ttctqqqqta	aataagttgg	tgcaaccggt	47340
actgaaaact	atttactaat	atctactaaa	accgaggaga	tgcacagact	acaaccaagc	47400
						47460
				ctcactcaac		
				aaatgcaaac		47520
caatcaaagt	caaatgtata	tctatattag	ggatatatac	aatggcatat	acacagcaat	47580
				gtaatctcag		47640
gagtaagaa	gggagatgag	ttgaggtgag	aaatttgaga	ctacctooc	caacacggtt	47700
gggtaaggta	ggcagaccac	ccgaggccag	aaacccgaga	ctagcctggc	caacacggcc	47760
aaaacctgtc	cccactaaaa	acacaaaaat	tageegggea	tagtggttgc	aggeetgtaa	
ttccagctac	tcgggaggct	gggttgggag	aatcgtttga	acccgaaagc	cggaggtcgc	47820
				tagagcaaga		47880
				atccgcaaaa		47940
						48000
				caaagatttc		
ggatągctac	tatcaaaaaa	agagagaa	taacagatgt	tagcaaggat	gtatggaaac	48060
tgaaattctc	acgcattgct	ggtgagaata	taaaatggtt	cagcetetge	ggaaaacact	48120
				gatccaacaa		48180
gggtatatata	CCanatasct	da a a de a ca e ca	tettesace	atatttgtag	acceptoste	48240
gggtatatat			terryaagag	atatttgtac	tttastast	48300
acggcagcat	Lactcataat	agetatgatg	Lggaaccaac	ataaatatcc	LCLYacadat	40300
						•

```
atatggataa gcaaaatgtg gtgtatacat tcaatggaat attaattagc aataaaaatg aagaaaattc tgacacatgc tacaacatgg atgaaccttg agggcattac attaaatgaa
                                                                                        48360
                                                                                        48420
                                                                                        48480
ataagccagt tataaaaaga caaatactat atgaggtact atattagata ctcatgcaag
                                                                                        48540
qtacctaaaa taggcaaatt catagagaca aaaagcagaa tggtggttgc caggggctgc
ggtaatggat acagagette aattttgtaa gatgaaaaaa ttetggagat tggttgcata
                                                                                        48600
acaatgtgca cacacttaac actggggaac tgtaaactta aaagtagtaa atggtaaaaa
                                                                                        48660
taaaaataat aaataataaa ttttatgtta ttttaccaca atatttatta aaagacaaaag
                                                                                        48720
                                                                                        48780
attaactaat taaacaaaat ccagccataa gctaatggta agagtaacaa ttaaagaaga
cacagaaaat tgaaaatcag tgactagaaa aagatattcc atataaatgc taacaaaaag
                                                                                        48840
caagtacagc aatataaaga gaatgaacaa aaaaaaaatt aaataagatg gctcgtttat
                                                                                        48900
toccaaaagg tacaattoac caagaagata caagaattgt gaacctttaa gcacataaaa cagettcaaa aatacaacat ttaaagaaaa atatatatta aacatagaaa tagtacaaaa
                                                                                        48960
                                                                                        49020
                                                                                        49080
acceptacaa gaateataat gggagtette aatacaacte tecatateaa caggteaaac
agagaaaaa aataagttaa ggatgcagaa aacctgaatt accatcaata aacttgagat
                                                                                        49140
taatatagaa ctgtataccc aatatactaa gagttcaggg aacagtcgtg actgacagtg
                                                                                        49200
gactgcaaat taatctgttc ttaatctttg tttttctttc agcactgtgg cagaatagag atcctaaaaa ccttccagct acaaaacatc tttttaaaaa tataaaaaaa tacaaaaata
                                                                                        49260
                                                                                        49320
                                                                                        49380
actotgaaat caatagaaga cacatggtga.aaccaaaatt ctagaataca gggagaataa
aggcattttc agatattaca aaaacagaaa attgatcatt gctgaagtaa tttctaaaga
                                                                                        49440
atgtacttga gggagaagaa aaatgttcca aagaaagta tctgtgatac aagaaggaat
                                                                                        49500
ggaaagtgaa gaaatggtaa acaggtagat aaagctaata aatgttgacc tagaaaataa
                                                                                        49560
caaaaacaat agcaataatg tetegttega aggettgaag taaaaataca attaaggeca
                                                                                        49620
                                                                                        49680
aatgtgaggt aagtggaatg aaagaattag aagteettge ettgtteaca ggaetgatta
aataaatgag ccaggttttc cattcaaaca gttaaaactt gaacaaaata aactcaaatt aagtagaaag ataaaaaaca gaaattaatg tcatagaaaa ataaaaaaatc aatagaatta
                                                                                        49740
                                                                                        49800
atcaataaat cctggttaat aaaagctggt tctttgaaag gattaataaa ataatcatta agcaagtctg atcaaaaaa aagagaaaag gtaccaaaaa aagtactgta tcagaaagag
                                                                                        49860
                                                                                        49920
                                                                                        49980
aacatacaga tacatacaga tatgtaagag tetgttttet tacaccagaa tactatatac
aacattatgc tagcatatat taaatttcaa taatgttaat gattttctag gaaaacagaa aatattaaat ttactttgaa gaaacagaaa aactgagaaa aataaatgat catgaaaaaa
                                                                                        50040
                                                                                        50100
atgaaaaggt aattaaatac tgatattaac tgcctaaaca acaccagcag cagcccaggc agtctgcagt caagttctgc caaacttgag ggaacagata attcttctat tccagagcat agaaaatgat ggaaagtttc ccaatttaat cagagaggac agcctgatcc ttgttatgaa
                                                                                        50160
                                                                                        50220
                                                                                        50280
cacagataaa aatggggtaa actatatgcc aaactcagat accaaaaccc taaataagat
                                                                                        50340
                                                                                        50400
gctagcttat tgatgtgaac aatccaaaag tgcattttaa attagcccag ggttttagag
                                                                                        50460
aaagaaaatc tagcaatgtg accaccactt atgttaacaa ttttaagacg aaaatctaca tgatcatatc aatgcatgct acacaaaagc atttgggcaa aaaacccaac acccacctt
                                                                                        50520
gactttttaa actcttagta attaggcata aacagaaatg tacttaatgt gatagaatac
                                                                                        50580
actoggtgaa gatacagagg gaatgotoco taaaaccaag cocaagacaa agattootat ttaacotoaa tagtoaacao tgcagogaga gtaatotatg gaagacaagg aaaaaagtaa aaacatgaga gacatotgtt gtttaacaga caataagato acotacttgg aagaggcaaa
                                                                                        50640
                                                                                        50700
                                                                                        50760
                                                                                        50820
cgaatcaagc gaaaaactat taaaactgag acaggcttta gtatggaggc tcagcttcag
ctgtagtttg ggctaccaaa ttcaactcgc ttgcttggag agttaatcct gcaaagctaa
                                                                                        50880
tttctgttga ggtattagga ttgaacagcc tgtgctcctc cctcctcccc catcttcaac actgaaataa cacggtgttt ggaactggat aacagaatct tccaaaaaca aaaattgtcc
                                                                                        50940
                                                                                        51000
                                                                                        51060
tgaagggotg acttgtgccc ttactcaaaa aacactttat ctgctgcctg cagctcctac
agttgctggt ggataagcct gccaaccagc tcggcgtaat tcttcctgca gagggcaagg
                                                                                        51120
aagagcactt tcacaggaaa atttttttcc gaactgtatg ccgcttatta cataaactta
                                                                                        51180
cgtgctggca aatggagctc cagcaaaata agatattcag agtcaaactt ccttaggaaa
                                                                                        51240
aaaaaaaaa aaaagcaagc acataacact aatttccttg catgggcact ggggaaggag
                                                                                        51300
gtogttactt cogcaegece geaggteege accaeeggga aacceaeggg caeegegege
                                                                                        51360
                                                                                        51420
tgeeceeggg cetteeaggt geactgegee geggegeeee agetgaeeeg ggatgegeag
coctagoeot teceetgica ecceggedag gaaggggegg gagegeggeg gaegeegagg
gegaaggget teteggteet etgeaceaeg cageaceee aaggeacaae agggagggtg
                                                                                        51480
                                                                                        51540
                                                                                        51600
cgggaggete ccgagaccca ggagccgggg ccgggcgtgc ccgcgcacct gtcccactgc
                                                                                        51660
ggcgaggget ggggtegeet eeagggeege agetgteggg agecaeetgg eteteagtee
                                                                                        51720
cgggtccctg cgacaacct cgggcccgga ggggaggagg cggccacctg ccgctgccac
ctgeggèace ggtcccaccg ctccgggccg ggcaggacag gccaggacgt ccctcctggg
                                                                                        51780
ctggggacag gacacgcgac gaggggaccg gggcccccgc ggcgaagacg cagcacgcct
tcccagaaag gcagtcccgt gcccccacga cggactgccg gacccccgcg ctcgcccgcc
                                                                                        51840
                                                                                         51900
```

WO 01/27857

```
51960
catecettea gaccaegegg etgaggegea aagageegge eggegggegg getggeggeg
cggctagtac tcaccggccc cgctggctca gcgccgccgc aacccccagc ggccacggct
                                                                                52020
cegggegete actgatgete aggagaggga ceeggegetee geeggegeet ceagecateg
                                                                                52080
                                                                                52140
ccgccagggg gcgagcgcga gccgcgggg gctcgctggg agatgtagta cccggaccgc
                                                                                52200
egectgegee greeteette ageeggegge egggggeeee eteteteeea geteteagtg
totcatotoc otatotgoto atoctotogot ogoacataat ogatottogo gogtoccaag
                                                                                52260
ccagatgtgg accccattte egeactetae actggaggtt ttetaagggt ggtgeeegga
                                                                                52320
ccagcagett cagceteate tgggaacttg agaaaatgca gatteteegt cccacceage etatteggtt ttteetgcae taaaaccatg aaggtgggge ccagcagtee acattetege
                                                                                52380
                                                                                52440
aagecegtea agtgattetg aggegeeete eagtttgaga getatgetea eggeeteace
                                                                                52500
tecgeceege aaggageeeg gtettgeetg tggegetage egeacaegga caceteatee tgeggggeee geeeeeege tgeacetea eegeecaaeg eeteeteegg gatgeagegg
                                                                                52560
                                                                                52620
                                                                                52680
aggogoctgg aagtoggoaa ggtoaacato cocotoagoa tottocotae cotoacggot
cetectecag gggtgeetea tggccagggg ttagaaagag ceaetgtgtt tettgacatg
                                                                                52740
                                                                                52800
gaagtggcct aagaccttaa tgaaaactgc aggagtggaa tgacagaacc tttggtcata
cttgagggcg tgaagctcaa atgaggagga aggaaaggat ccagggagaa taaccaaccc
tggcaagttg tggcgcccag gtagagggc gagcctaggc tagcggttct cgaccagggc
                                                                                52860
                                                                                52920
cggtgttgcc cctcctcgcc gccccgcgta catttgggga ggtctggaga catttttggt
                                                                                52980
                                                                                53040
tgtcatgatg cgggagttgc tactgttgcc taagtgggta gacacgaggg tgctcctcaa
catcctacct gaaggacagg actgccccac aaggaagaat gatccggccc caaataagaa
                                                                                53100
accetggget ggteageaac aacceetttg ttetgagaag agaggaggaa agaataaaag
                                                                                53160
aagtggggtg aagttitggt ttggtagagg aaacttgaag acattitcac tggaaaggaa
                                                                                53220
                                                                                53280
gagaggaaga ggagggagat gtetgtaagg aegageaaae egggtgaeag etgattteet
                                                                                53340
catattgaag taatgagtoo tagttataat aaattootaa taaaaaccca gtttatooot
gcaataaact tgtcttttt ttttaaatat actgcttgat tctgtttgct aatattttat
                                                                                53400
ttacaggett tgcattgata tgcaaaaatg agatgggcaa taattttett tttgaatgte taatgttgtt tggttteaga atcaatgtta tgcteacate ataaaaaatt tggaacegag
                                                                                53460
                                                                                53520
gcaggaggag tgcttgaggc cagaagttcg agaccagtct aggaaacaca gtgagacccc cccatctcta caaaaaaaa aaaagaaaaa aaaatgggca tgtttgcttt ttccttttac
                                                                                53580
                                                                                53640
                                                                                53700
tctgaacaat ttaaggagca ttaaaattat ctattctttg aggtttgatc atttcccagt
taaaaatgtt cctccagcc tgatgctttc tttggggagg gtaaatcttt taaggctaga aaagtttctt ctgtggcaat tttattattt acattttaaa aattattcta gagttaattt
                                                                                53760
                                                                                53820
tgataaagca tgtatttett aaaacaaatt ateetttttt teeagatgtt caagtgtatt
                                                                                53880
tgcataaagt tgaggaaagt agtcttttgt gaatctttta acttctccca aatatcttat
                                                                                53940
tttgtgtatt tttgcttctt tattttgtta acttttaaaa gtgtatttt ttttcaaaga atcagctctt aggtttatgt ttttggttat actggagctt ttttcttctt ctttttaaaa
                                                                                54000
                                                                                54060
tattttttct cctttatttt ttagacgtat tttgatctaa cgtaatcgga agaaggtaaa
                                                                                54120
                                                                                54180
ttagaatett tigitaetat igigititta titeteetta titeteigaa gieeigetti
                                                                                54240
ataaatagta ccatgttatt tgtgcataaa tattcatttg tcttatattc ttgggaattt
toccactica toataaaatg accttoottg totcatttaa tgtgttcaaa ctttgccctg
                                                                                54300
aatttaactt tgtctgatat tttaccatcc tgctgaattt tgtttgttac cccaaacaac ctttgctgtt ttcgtcttt ctgaaccett tattttaggt aatcccttga attagagcac
                                                                                54360
                                                                                54420
                                                                                54480
taagttttgc tttgtgatta aatctgaaaa tctttatctt gccatagatg agttgagccc
taticatgig acagetatat tatgetgttt catagecett itggteettt titcaetett
                                                                                54540
gcattgcata ttttgtgttt attgtgtttt gtgtttcttc tgataatttg gaaggtttgt
                                                                                54600
atttttatte agggagttge ettataatea taeteegeaa taeacategt eeteagttte
                                                                                54660
ttcagactgt ctgttaactc cctattctga ataaaaatga cattgtaatt tccctctttt
                                                                                54720
                                                                                54780
ttotitacec ctittettet ceteacetaa tgtaaatgat tttateette tttagtattt
gettttttaa ttaactacat ttataaatat etttateaet tgatttttaa atcagetttg
                                                                                54840
aatgagatat ttggattoot agatataaaa gatgttaatt ataccattto cacgttagta
                                                                                54900
ggtttataaa atcatacatt ctgctgtgta accataatcc cacgtttgtt ttagttccac tcctacagtt aaaagattca gaagtattat taacagttat tttgccatag tttttcccc
                                                                                54960
                                                                                55020
                                                                                55080
aacccatttt gtggtaagtt atgateetge tttagtttet taagaataat ttatagagea
                                                                                55140
gagtgtggtg gctcacgttt gtaatcccag cactttggga gacaagaggt agaaggatcg
cttgaagcca gcagttcaag accacctga gcaacatagt gagaccttgt ctctacaaaa
                                                                                5,5200
                                                                                55260
aattttaaaa tttagccaga cgtagtggcg tgtgcctata gtcccagcta ctcaggaggc
tgaggcaaga ggattgctag agcccagaag titgaggctg cagtgacctc tgattgtgcc
                                                                                55320
                                                                                55380
actgcaèccc agtctgggca agaaagtgag aacctatctc tttaaaataa caataataac
ttatgaaaat tatatteeet gagttittea tgtttaaaaa tatttgttge etttateetg
                                                                                55440
taaaagtttg agtataaatt ctigggttat actttattta ttgaagaatg tataagtatt
                                                                                55500
```

55560 gtottotaga attgagtgtt gotgtaatga aaccagaagt cagootggtt tatttttoot cagaaatgag gtaattgeeg geeggacaee gtggeteatg cetgtaatee caacaetttg 55620 ggaggccgag acaggtggat cacgaggtca ggagattgag accatectgg ctaacatggt 55680 55740 gaaaccccgg ctctactaaa agtacaaaaa gttagctggg catggtggtg gacgcctgta 55800 atcccagcta cccgggaggc tgaggcagga gaatggcgtg aacctgggag gaggagcttg cagagagetg agategegee actgeactee ageetgggeg acagagtgag acteegtete 55860 aaaaaaacaa aaaaaaaaca aagaagtgaa gtaattgcca tgatgctcca agaattatct 55920 ctttgtctat gaaatccaga aatctcactg ttatacattt tggaattatt attctgggcc aatatttcct gggacacaat agattgactc tatagattta attttttt ttttttgag 55980 56040 acagagtete actgeaatet cagettaetg caacetetge etcaegggtt caageaatte 56100 tectgeetea geeteecaag tagetgggae tacaggegeg tggeaceatg cetggetaat ttttgtettt ttagtagaga cagggtttea ceatgttgge caggetggte ttgaaegeet 56160 56220 aacctcaagt gatccacctg cctcagcctc ccaaagtgct gggattacag gcgtgagcca ccatgccag cctcaattcc tctttctatc tggtaatttt tctgaagttg aaaacatttg 56280 56340 ttotaatacg ttatttcagt gttottotaa gatgtgtaaa gcaccctatt cccaggtcag 56400 ccccatctt gctagtgage teggetggtt etteacaaga getetggttt teteetgett aateteaagt acctetgtea geeteeacet ggtttatgat teggagtttt tetggttttg 56460 56520 ttttttgttt ttgacagagt cttactctgt cacccagget ggagagcagt ggcataatct 56580 cageteactg caacetetgt eteccagget tgagegatte teetgeetea geetactgag 56640 tagetgggat taeaggegeg tgecaccaea eceggetaat ttttgtattt ttagtagaga 56700 tggggtttca ccatgttggc cagggtggtc ttgaactcct gacctcaggt aatccacctg 56760 cctcagcctc ccaaagtgct gagattacag gcgtgagcca ccgcgcctgg catggtttgg 56820 56880 agttttaatc tgtagtttta ataaagatag tgcttatgtt tgtgtttctt atatttcttg 56940 57000 gotgtgtcae ttetggagge tggagtgeag tggegegate teaggtcaet geaaecteeg tetecegggt teaageaatt etectgeete ageeteeega gtagetggga ttacaggeae 57060 57120 57180 ctgccactit ttaatttttt tagagacaga gtctcgcttt gttgaccagg ctggagtgcg 57240 gtggtgcaat catggctgac tataacctcc aaatcctggg ctcaagtgat cctcctgcct cagectectg agtagetggg actacaggea catgecacca tgeccagtta attttaattt 57300 ttttgtagag acagggtctc catatgttgc ccaggetggc etectactce tggcetcaag taateeteet aceteageet eccaaattae taggattata ageatgagee aceatgeeca 57360 57420 57480 gectigitet actaettiaa titeatatgi taggigaeea tgiaatigat eateeaaaee aggatactgt aagaatgaaa gaggetgaca gtagtatgat getgggaeta geattgtgea 57540 ctgagattat ttctgggaaa gcaggagata cggtcaccct acttatagtg tgcttgtctt 57600 57660 tggattgttg aatttggagt ttctatttgc aggcttattt caactgggca gccttgatcc gecetgeeca geaatgetae egttetetee acegggtete tgggacceet teagteacta 57720 tacttagete agttececae ecteceaete ectanaageg taaccaggaa teetgeetea 57780 ggtctactgc cgtcttccgt gggctgtttc agttcctatt acccagagtc aaactcccag cattccctac ctgattccag acttggagtc cagagcttta acctcttcag gccaactccc cactttgcat ttctgtccct atatcttagt ccatggagat acatttcatg tctttgagtc 57840 57900 57960 tacttacaaa gtaaattttg ctgtttttta atttttttt tgagatggag tcttgccctg 58020 58080 tcacccagge tgtggtgcaa tgacgccate teggeteaet gcaaceteeg cetectgggt tcaagcgatt catctgcctc agcctcccaa gtagctgtga ttacagacag gcaccaccac 58140 gcccagctaa tttttttat cttttagtag agacagggtt tcaccatgtt ggccaggctg gtcttgaatt cctgacctcg tgatctgccc atctcggcct cccaaagtgc tgagattaca 58200 58260 ggcgtgagcc actgtgccca gccaatittg cttttitat atttcaitgc tatatgttta 58320 gaggataagt ttacagtgct atatgcattc ccaaatatta gaccaaaaaa atctccaaaa 58380 aattagaaag aaaatccaaa aaatctcaaa aaataccaaa aagcaacaat ctcacagacc 58440 atactcactg acccccaata aaataaaatt agaaattaac cacaacttaa caaaataaag 58500 tactcaagtc agagaggaaa gaggaaataa acatcaaaat tacaaagtct aggcggtggc 58560 tcacgcctgt aatcccagca ctttgggagg ccaaggcggg cagatcacaa ggtcaggaat 58620 tegagaccag cetggecaat atggtgaaac ceegttteca etaaaaatac aaaaattage 58680 58740 caggcatagt gatgtgtgcc tgtaatccag ccacttggga ggctgaggca ggagaatcac tgaacccagg gagacgaaga ttgcagtgag ccaaaatcgt gccactgcac ttcggcctgg 58800 gtgacaaage gagactecat etcaaaaaaa aaaaaattae aaactettta gatagaaatt ttggtgtttt tttttgagae ggagteteae tetgtegeag aggetggagt geagtggae 58860 58920 58980 tatgtcagct caccgcaacc tccatctcct ggattcaagc aattctcctg tctcagcctc ccaagtaget aggattacag gegeceacea ccagaeceag etagttttta tatttttagt agagatggtg tttcaccatg ttggecagge tggtetcaaa etectgaect caagtgatee 59040 59100

```
acctgcttca gcctcccaaa gtgctcagat tacaggcgtg agccaccgca ccccacctag atagaaattt caacatgagg ccgggcacaa tggctcacgc ctgtaatctc agcacttcag
                                                                                        59160
                                                                                        59220
                                                                                        59280
gaggetgagg egtgggagga teaettggge eeaggagtte aggaceagea tgggtgacag
59340
                                                                                        59400
ttcaacatga aaagtatete teaaaceett egagatgttg geaaaaageg aeteaaagga
aaatgtatta ctgtgtgtga atttgcttga aaataagaaa gaggccgggt gtggtggcta
                                                                                        59460
                                                                                        59520
acacetgtaa teecaacact etgggagtee gaateaagtg gateatgagg teaggagate
                                                                                        59580
qaqaccatcc tggctaacat ggtgaaaccc tgtctctact aaaaatacaa aaaattagct
aggegeggtg geleatgeet glaateeeag eactttggga ggetgaggea ggtggateae
                                                                                        59640
ctgaggtcag gggtttgaga ccagcctggc ctacatggtg aaacctcgtc tcttctacaa
                                                                                        59700
                                                                                        59760
atacaaaaat tagetgggeg tggtgggg tgeetgtaat eecagetact cagaggetga
                                                                                        59820
ggcaggagaa tegettgaac eegggaggeg gaggttgegg tgageegaga tegeaceact
acactecage etgggeaaca geetgggtga cacagtgaga etccatetea aaaaatacaa
                                                                                        59880
aaaattaget gggtgtggtg geetgegeet gtagteecag etaceeggga ggetgaggea ggagaatgga gtgaacetgg gaggaggage ttgcagtgag eegagateec accaetgeae
                                                                                        59940
                                                                                        60000
tccagcctgg gcgacagagc aagactcttg tctcaaaaaa aagaaaaaa aaggaaaaaa gaaccctgat aataaagaaa ccaaatgttc aactctcaaa gctcggacac tttaaagaaa
                                                                                        60060
                                                                                        60120
taattaataa aggcagaagt taaagggagg atgataaagc aattititt gtiggtitti
                                                                                        60180
ttgagatgga gtottgotot gtoaccoagg otggagtgca gtgatgogat ottggotoac tgcaacctot gcotocoggg ttcaagcaat totcotgcot cagcotoctg agtagetggt
                                                                                        60240
                                                                                        60300
actacaggtg cgcgccacct ggcccagcta atttttgtat ttttattaga gacggggttt
                                                                                        60360
caccatattt gttaggetgg teteaaacte etgateteag gtaatetgee caccteggee teteaaagtg etgggattae aggeaggege cacegegeet ggeetaaage aaaatattgg
                                                                                        60420
                                                                                        60480
ttotgtgcaa aaggtcaata aaaagagcaa acgtttacaa actggagcca gcacccattc
                                                                                        60540
ageteagtgt gtetggagaa aaaacaatet egetteagaa tteatgatta egeageeett
                                                                                        60600
                                                                                        60660
titgetiect aaaaateeta etatgttget gttgaceatt etetetett etetetete
tgetttetet ceagaaaage tatteagaca tteteetett teeteaaace teeaacaett ceteeteat cettageete agetgetgae eteaetteta ateattgaga aaccaggaga
                                                                                        60720
                                                                                        60780
agcatttaag agtgaacctc cgcctccccg cacgggcaaa accacccacc cacagaattg tgccccaatt ctgcgtcctc tcctctcacc atggatggac ggtccaggct ccgagccaaa
                                                                                        60840
                                                                                        60900
                                                                                        60960
gccaggeete ecetggaget etggateeae cacetgeage tteteaggea gggeeecage
agetecectg etecettgta ceateaatee etececteae tgggteacte ceaacaatat
                                                                                        61020
                                                                                        61080
atatattag tgatgttict cccatgtggt aaaatcactt agectetete etececeage
tactatecta titigittett tecattetet geaaaaette teaaageatt gtgtetatgt getgaeteea titatettet eeegttetet getgagteet teeeaeagae teteaeeeea
                                                                                        61140
                                                                                        61200
gttactccat gaaatgacct ctgcactgcc acatccaatg gtgaatgttc agttcttaat tttattcagt ctttcagcag catttgacct ggccgatcac tccctcttct taaaaatact
                                                                                        61260
                                                                                        61320
tttetcagec aggegtgatg getcacacet gtaateecaa caetttggga ggecaaggeg
                                                                                        61380
ggaggateat gagageeeag gagtteaaga teageetggg caacatggea agaeeetate
tetacaaaaa etaaaaagta geeagtgtga tggeatgeae etgtagteee atetaettag
                                                                                        61440
                                                                                        61500
gaggetgagg cagtaggatg acttgageet gggaaateaa ggetgeagtg agecatgatt
                                                                                        61560
gcaccactgc actecagect gagtgacage gagaccetgt etcaaaaaga caaaatagga aacttttete agcatattee tetgattete etgetgette tgtetgeaca gatteagtet
                                                                                        61620
                                                                                        61680
cettigeegg tietteetea teeteetgat etetigaeet igaagtgeee eagagtaeag teittitti tittitigag aegeagtete gietgieaee eaageiggag igeaatggeg
                                                                                        61740
                                                                                        61800
                                                                                        61860
aggteteage teatgeaace tetgeeteet gggtteaage gatteteetg ceteageete
ccaagtagcc aggactacag gcacatgcca ccatgcccag caaattgttg tatttttagt
                                                                                        61920
agagacaggg tittactata itggccacgc tggtctcaaa ctcctgaact cgtgaaccac
                                                                                         61980
ccgcctcggc ctcccaaagt gctgagatta caggcatgag ccaccacacc cggcccagag tacagtcttt agacggcctc tctacctata cttgctcccc tcataaactc ctcctgcctc
                                                                                        62040
                                                                                        62100
atggetttaa ataccategg tagaetgatg acteceatat ttetettttt tttttggaga
                                                                                        62160
cggagteteg cteagteece caggetggag tgeagtggeg cgatetegge teactgeaag etceacetge caagtteaca ccatteteet aceteageet etceagtage tgggactaca
                                                                                        62220
                                                                                         62280
                                                                                        62340
ggcaccegee accaegeetg getaattttt ttgtattttt agtagagatg gggttteace
atgttageca ggatggtete gateteetga eetegtgate egeceatete ggeeteeeaa
                                                                                        62400
agtgctggga ttataggtgt gagccaccgt gcccagccga tgactcccat atttctatct
                                                                                        62460
cttgctgtgt gggagttete etcagaacte catacteata aatecaacte teataaatag
                                                                                        62520
tateteaaat gggeaatatg eteaaaagte aatteetaet ttteteeeta aaettgettt
                                                                                         62580
cctgcagtct ccaccatctt aatgtccaat ctaacattag gaggcaaaaa ctttgaagtc attcttgact cttctctatt acacacccta tccaatcttt ctgcagatcc agtcgaccc
                                                                                        62640
                                                                                        62700
```

```
caaatccagt tageteteat cateteceet gttaceceet ggteeaggee atetteetet
                                                                                62760
ctcacctgaa tcactgcagc attctcctca ctggtctctt tggttctgtt ttcactccac
                                                                                62820
                                                                                62880
cttagcatag totocacaga gcagtcagag ggatcotttt aaagtgtaat toccatootg
tccctgctct gctcaaaacc ctgtcgtgat tcccgtttta atctgtcaga ttaaaagcca
                                                                                62940
gagtetttee agtgacetae atgatetgee tattateace teccaettet tteccettge
                                                                                63000
                                                                                63060
tcactccact ccagctetge agetgteett tetgttteet gaacageeca gattttgett
                                                                                63120
ctttagaacc tttgtatttg ctgtcccctc tgtctggaat gtttttccag gaagtcacct
                                                                                63180
qqctctctcc tgcacttcct tcctgaccac catgtttaaa aatcactcaa acacacttca
ggccggacat ggtggctcac gcctgtaatc ccagcacttt gggaggccaa ggtgggtgga
tcacctgagg tcaggagttc gagaccagcc tggccaacat ggtgaaactt cgtctctact
                                                                                63240
                                                                                63300
                                                                                63360
acaaatacaa atagtagcca ggtgtagtgg cacacacctg taatctcagc tactcaggag
getgaggeag gagaateget tgaacceaga aggeagagga ggtgeagtga gecaagatea egceacaaca ecceageetg ggtgacagag caagacecea tetcaaaaaa aaaaaaagaa
                                                                                63420
                                                                                63480
aaaaaaatca cacaaacaca cttctcttca tattcctttt ccaagtttta tttttctcca
                                                                                63540
gaatacttta cattgtttta atggaagttc tccgtttccc cccaactaga atggatactt
                                                                                63600
                                                                                63660
cctgcaggta ggcactctag tecteccate caagtactaa ccaggetcaa ecctgettag
cttctgagag caggggagat caggcctgtt cagggtggta tggcccagga attttgattc
                                                                                63720
tgttttatte attgetgtte tgttgattet ettttgttee teeteetagt getgagaaca
                                                                                63780
ctacttgtac ataataagca ttcaataaat atttgttgaa tgaatgactt gttgaatgaa
                                                                                63840
ttaatctcag aaatgcagga ctggttctac attagaaaat ttttcaaggt cattctctgt
                                                                                63900
tgtcgtaaca cattaagaga ggaaaatttt gtactctaaa tcatttgata aaatacatac
                                                                                63960
                                                                                64020
tgatttctgt tttcaaaaac tcttagtggc tgggcgaggt ggctcacatc tataatccca
gcattttggg aggacgaggt gggcggatca cttgaggtca ggagtttgag accagcctgg
                                                                                64080
ccatcatggt gaaaccctat ctctactgaa aatagaaaaa ttagccgggt gtggtggcgc
                                                                                64140
atgeetgtag teccagetae etgggagget gaggeaggag aatggettga accegggagg
                                                                                64200
                                                                                64260
cggaggttgc agtgagccaa gatcatgcca ttgcactcca gcctgggtaa cagagtgaga
ctccatctca aaagaaaact cttagtgagt ttaggaatcc aaggaagacc ctcaaactaa
                                                                                64320
atagataatc tagctaccag aagccttcag taaaccttaa cactccatgg tgaaacatta
                                                                                64380
gaaacattcc tactaaaaga caggctaaga atgcctgcaa tcttcacggc tagtccaaga
                                                                                64440
                                                                                64500
agtcaaaaag aagaaatgag cgctgattta aaaaaataaa caaacaaaaa actaccgatg
cagaggetgg cagcaaggac tgaaggactg tacagtactt gcctggagca ggcggatggc
                                                                                64560
cacaccctg cgaagcctgc tcagctggct gggggacgct ccagtgtgtg agtggcagga tgcagggtac ttcctctgcc agggagttgc actggggaga tcctccccca ctcacacttt
                                                                                64620
                                                                                64680
ggcagctggg gctttggaat gtgacttagc ttctgtcaaa gggtcaatcc accetttgat
                                                                                64740
                                                                                64800
atatgatgca aaggcgaaca tatgatgcaa aggtgagaga acagcccaaa ttaggacttt
                                                                                64860
taccacaget gtggaggtgg acagegacag tggtgggeec tggecagaet tttcatgete
aaaggtggtg gttgttcttc ctacttcttg tccctccagg gcttcctttg cctgtgtgct gaacctgctt cttttaattt tttttaactt ttttaaattt ttaattgttt taattaaaac
                                                                                64920
                                                                                64980
aaatttīgaa aactgtotga acctgotttt gaaccotgot atgattīgaa tgtttgtooc
otgocaaact gattttgaaa ottaatotoo aaagtggoaa tattgagatg gggotttaag
                                                                                65040
                                                                                65100
                                                                                 65160
cagtgactgg atcatgagag ctctgacctc atgagtggat taatggatta atgagttgtc
atgggagtgg catcagtggc thtataagag gaagaattaa gacctgagct agcatggtcg cccettcacc atttgatatc ttacactgcc taggggctct gcagagagtc cccaccaaca
                                                                                65220
                                                                                 65280
agaaggetet caccagatac ageteeteaa cettgtaett eteageetet gtaactgtaa
                                                                                 65340
                                                                                 65400
gaaataaatg ccttticttt aigaattacc cagtitcaga tattcigtta taaacaatag
                                                                                 65460
aaaacgaact aaggcaaact ctcatgattc tactgccatg ccattccaat aaactccctt
tatgettaag agagecagag ttggecagge gtggtgaete aegeetgtaa tteeageaet ttgggaggee gaggeaggtg gateacaagg teaggagate gagaecatee tggetaaeae
                                                                                 65520
                                                                                 65580
ggtgaaaccc cgtctctact aaaaatacaa aaaaattagc tgggcgtggt agtgggtgcc
                                                                                 65640
tgtagtccca gctactcggg aggctgaagc aggaggagaa tggcgtggac ccaggaggcg
                                                                                65700
                                                                                 65760
gagettgeag tgagtegaga tegtgecaet geactecage etgggtgaca gaatgagaet
ccgtctcaaa aaaaaagaga gccagagttt atttctgttg cttgcaacca agaaatctgg
                                                                                 65820
ctggtgcact gaagtttcca taaataatag caatttaaag actctttcca agccaggcaa
                                                                                 65880
                                                                                 65940
tgcctagcct tgtgtagtcc ttgtggtaat acattcattc attcatttgt tcaaccaact
                                                                                 66000
gtgctccaga gactaagaat acaaaaatgg gggccgggtg tggtggctca cacctataat
cctagcactt tgggaggccg aggcaggtag atcacctgag gtcaggagtt cgagaccaac ctggccaaaa tggtgaaacc cctactctac taaaaataca aaaaattagc tgggggtggt
                                                                                 66060
                                                                                 66120
                                                                                 66180
ggcggačacc tgtaatccca gctactcgtg agactgaggc aggagaatca cttgaacccg
ggaggcagag gttgcagtga gccgagatcg caccactgca ctccagcctg ggcaacaaga
gcgaaactcc acctcgaaaa aaaaaaaaa aaaaaaagag ggccggggct gggcgcagtg
                                                                                 66240
                                                                                 66300
```

```
gctcacgcct gtaatcccag cactctggga ggccaaggca ggagaattac gaggtcagca
                                                                              66360
gategagace ageetgacea acatggtgaa acceeatete tactaaaaaat acaaaaatta
                                                                              66420
                                                                              66480
teegggegtg gtggegeaca cetetagtee cagetaettg ggaggetgag geaggagaat
cgcttgaacc cgggaggcag aggttgcagt gagccgaaat catgccactg cactccagcc
                                                                              66540
                                                                              66600
tgggtgacag agtgagacto cgtotcaaaa aaaaaataaa aaaaaaaaaa gaattcaaaa
attgtagagt tatagtgtgc tictagttta gttgagagga catctgtcct tcaaggaagg
                                                                              66660
                                                                              66720
ctagaatcta taccetgagt cettactgaa atcaatccag cagtcaaaac atgggaccaa
                                                                              66780
cgatcacagc agtaagatag gaagagcacc tttgtacatt tagctcatgt tgagataagc
                                                                              66840
cactgacaga gctgaaggaa gctcacagtt ctgggttcca tcctttggca tttaaaaaga
aaagtgctaa gaaaattegg ttggteaegg tggeteaege etgtaateee aacaetttga
                                                                              66900
gaggecaagg caggeagate acgaggteag gagttegaaa ceageetgge caacatggtg aaaceeegte tetactaaaa acagaaaaat tageegggea tggtggegea tgcetataat
                                                                              66960
                                                                              67020
cccagctact caggaggctg aggcaggaga attgcttgaa cccgggaggg ggaggttgca
                                                                              67080
gegagtgaga geaggeeact geacteeage etgggagaea gageaagaet etgteteaaa
                                                                              67140
                                                                              67200
ttcaggccag gccaggcctg gtggctcaca cctgtaatcc caacactttg ggaggctgaa
                                                                              67260
gegagaeggt geettageee aggagtttga gaccageetg ageaacatag egagaeeetg
                                                                              67320
tetetataaa aaaaaatttt tttttggeea gaegeagtgg etcaegeetg taateeeage
                                                                              67380
actttgggag gccgaggcag gtggatcacg aggtcaggag atggagacca tcctggctaa cacggtgaaa ccccatctct actaaaaaat acaaaaaatt aaccgggcgt ggtggcgggc
                                                                              67440
                                                                              67500
gcctgtagtc ccagctactc gggaggctga ggcaggagaa tggcgtgaac ccgggaggcg
                                                                              67560
                                                                              67620
gagettgeag tgageegaga ttgegeeact geacteeaga etgggagaga gtgagaetee
gtctcaaaaa aaaaaaaaaa aaaaaaaaat taattgtcag gtgtgctggc atgcagctgt
                                                                              67680
agtectaget actegggagg etgaggtaag aagategett gageecagga gtteaagget geagtaatag tgeeteteae tetaceetgg gtgacaatga gaeeetetet caaaaagaaa
                                                                              67740
                                                                              67800
                                                                              67860
67920
aaaagaaaag gaaggaagga agaagaaaaa aaaagaaaga aagaaaagag agagaagttc
aaagaccaaa gggtcaggat cccaaaatag tttttatgtt ttatttattt atttacttat
                                                                              67980
ttattttga gacagtatgg ctctgtcgcc caggctggag tgcagtgatg cgattgcggc tcactgcagc ctccaaactg ggctcaggtg gccctcccac ctcagcctcc cgagtagctg
                                                                              68040
                                                                              68100
                                                                              68160
ggaccacagg cgcgtgccac catgcccagc taatttttta attctttgta gagatgaggt
ctctatatgc tgcccaggct ggtctcgagc tcctgggctt aagccatcca cccgcctggg
                                                                              68220
cctcccaaag tgctgggatt acagaagtga gccaccgcgc ctaatcgggt ggtttgtttg
                                                                              68280
tttattgacg gggtctcgct gctgcccagg ctggagtgcc agtggctgtt cacaggtgca gtcctggagc attgcatcag ctcttgggct ctagcgatcc tccagagtag ctgcagctgg
                                                                              68340
                                                                              68400
gattccaggc gcgccaccgc gcggggctca gaatgggttt ttatattgag ggttatgctg
                                                                              68460
                                                                              68520
ccacctagag gatatatgta gtaccgaact gtgtgcgcag ggaggctgag gttgcagtga
gccaagatga tgccagggca ctccagcgtg ggtgacagag caagatttca tctcaaaaaa
                                                                              68580
aaaaaaaaaa aaaaaaaaa aagaattgaa agtaaggtot tgaagagata tttgtgcctg
                                                                              68640
                                                                              68700
tatggtcata gcagtattaa ctitgaccca ciagciaaaa cacaaaagca acatgtgtct
                                                                              68760
gtcagcaggt gaacggataa acaaaatgtg gtatatatgt acaattgaat attattcagc
ctttaaaaag gaataaaagg ctggatgcgg gggctcacgc ctgtaatcct aacactttgg gagactgagg tgggtggatc acccgaggtt aggagtttga gaacagcctg gccaacatgg
                                                                              68820
                                                                              68880
tgaaacttca tctctactaa aaatactaaa attagccggg catggtggca cttgtctgta atccaagcta ctggggaggc taaggcagga gaattgcttg aactcaggag ccggaggttg
                                                                              68940
                                                                              69000
                                                                              69060
cagtgagcta agatggcacc actgcactcc agcctgggca acagagtgag actccatctc
                                                                              69120
aaaacaaaca aacaaaaaat tattatttoc aaagaaacaa gaccctgggt ccatttccca
gcccacacct gatgttgact cacaacaca agcctggttt gctatgagcc tgcttcattt
                                                                              69180
aattgtcacc ttaacttcac atcaccetca agtectggaa taactetttg etgacetttg tgtgetgage catetecatg tegetcaacg tgcagtecet etcaetgeac tgagteaata
                                                                              69240
                                                                              69300
                                                                              69360
gecagaegtg gtetgaetge agggteatee ttggtggett aggetgaete gggeatagea
                                                                              69420
gggtgctctg agacctcacc gcatataggc tttgccccca ataaactcta tataatattc
                                                                              69480
atattatgtg gtctgggtgt gtgtagcttt gcactgtctt ctcgtgacag tgccctcaac
ctctttccca ggatttcctc ctctacctcc tcaagtccca ctgctctgca aagaccaaaa
                                                                              69540
getgeagagt cecageteec teetttacae eccaegaege ageeteetet eteagaacee
                                                                              69600
                                                                              69660
tttaaacaga gtottttact gcagatocca agaacagoca caccoototo toccacocac
                                                                              69720
tecaĝacaca eccaggiaat tatageacee agggiaacia igiagaigga gieceiggaa
catgtggata gtgccccctg ggagtatgca aaagcaacat tgctggcacc tgcagagaac
                                                                              69780
aggitgacat ccaggaatca gagcatgggc ctctgggagg tagggatgtg gccaggcagg ctgccaaaaa ttggtagagc aaggccacag gatctttctg accttccttc caaacagagg
                                                                              69840
                                                                              69900
```

WO 01/27857

```
ctcctgtact ggtgatccct gtgttgattg accaetecet teetgggggt egtggtetet
                                                                                           69960
gtcccagttg cccggacttc tgtgagtgtc ctactgaggt ccttttcatg agaagcatgc
                                                                                           70020
tgtccttcca cctgctggga gcaagagtga caacttcaat actataatag cagtggcata
                                                                                           70080
cagagaagaa gaaagatgaa gtggcaagaa aaacaggctt ccaagcagga gtttttctat
                                                                                           70140
aaaacaaaa acgtttacaa gcaaactttt tataaagggc tagatagtaa atattttagg ctttgagagc cacatagact tgtttgcagg gactcaatgt cgctattgta gtttgaaagc agccatcagg gttatgtaaa tgagtgagtc tgattttgtt tcagcaaaat tttatttacc
                                                                                           70200
                                                                                           70260
                                                                                           70320
                                                                                           70380
aaaacagaca atgagtggge tggatttgge ecatgateet tagtttgeea aeteetgett
                                                                                            70440
tgggeteace cagatetgat tttgaattet ggetetgeta etggttaget geaggagett
                                                                                            70500
ggaaggetet etgageetgt tteeteatet gtaaaattaa agcaataatt tetaacaete
agagtgtta ceteacget gtaateccag caetttggag getgaggeag geggateace tgaggteaga agtteaagae cagegtggee aaegtggeaa aaceetgtet etaetaaaaa atacaaaaag tageegggea tggtggegeg catetgtaat eecagetaet tgggaggetg
                                                                                           70560
                                                                                            70620
                                                                                           70680
                                                                                            70740
aggcagggat actgctagaa cctgggaggt ggagcgtgca gtgagtggag atcacacctc
cacactecag ectggeegae agagegagae tecateteaa aaaaaaaaa aaaaagagtg
                                                                                            70800
ttagaaggtt ttgagataat gaataaaaga tgccttgtgt atactaagta ttcaacaact
                                                                                            70860
                                                                                            70920
gatagetgea ttggtetaat tataacagtt tagaagegat tgagteaaca aatgetggat
                                                                                            70980
ttgtcaggga ggacttccta tcaggaggta gatcttgggc tgagtcctga agcaaagata
ggcattggat agaggagttg agagaacacc ctaggactgt tattattatt attcgacacg gagtetettg etetgteace caggetggag tgeagtggeg egatetegge teaetgeaac etetgeetee caggeteaag egatteteet geeteetaag tagetgagae taeaggtgtg
                                                                                            71040
                                                                                            71100
                                                                                            71160
                                                                                            71220
tgccaccaca cccggctaat ttttatattt ttagtagaga cagagtttca ccatgttggc
                                                                                           71280
catgctggtc tcgaactcct gacttcaggt gatccacccg cctcagcctc ccaaagtgct
ggaataacag atgtgageca cegcacecag cecagaacea tttttcaate ettggetetg
                                                                                            71340
cettttatta getgeaagat eteaggeaat ttatttaace tetecaaaga eteatttet catteacaaa atgaggeaaa taataatate taetateeca ggttgteatg agaattaaat geaacatgae atttaatgaa atgagaagte eettggacat taaetggeta aagtatgtge
                                                                                            71400
                                                                                            71460
                                                                                            71520
togacaagga tatcatttta ggtggatact tagcatotoa gaactgatgo toacaatgga atatcattga aacgcattaa aattcatttt aaatgattgt aggtagtgag gcaattgaaa
                                                                                            71580
                                                                                            71640
gaagaagaca agaggactga ttataatgct teaggeteae tagteteett ttaggaggga
                                                                                            71700
                                                                                            71760
aaaacaattt caagttaaat tttaggctct agatttttac ccctgctgct cattagaatc
acceagattg atgaaatcag ageccatetg aggetgtgtt tttcatetec agaatgagag ctgttgtggg gattaagttt ttgaaaaagt acatetaaca ggtgategaa aatgatagtg
                                                                                            71820
                                                                                            71880
atattattgc agtgatggtc attattgttg ttattattat actgaaagag gcttcagttt
                                                                                            71940
totgatocat aaagtgaggg aattgcatga gaccattgct aagatteett ctagetetgt
                                                                                            72000
                                                                                            72060
ttttttgttt ttgtttttta gacagagtct ctgtcgccca ggctggagtg caatggcatg
atcttggctc actgcaacct ccgcctcccg ggttcaaatg atcctctgt ctcagcctcc gaagtagctg ggactacagg cacaccacca catgcccagc taacttttat atttttaata gaggtggggt ttcaccatat tggtcaggct ggtctcaaac tcctgacctc aggtgatcca
                                                                                            72120
                                                                                            72180
                                                                                            72240
cccgcctcgg cctcccaaca tgctgggatt acaggcatga gccactgtgc ccaaccctt ctagctttct tgatcactga ttctagggtt ctctgctgaa atatattga gacatcctgg
                                                                                            72300
                                                                                            72360
                                                                                            72420
ataaaagatc atgcaagagc tcccaatatg gtattaataa ttgattctgg aggcttagct
                                                                                            72480
actectgatg gattagacat gacteaactg cetetettat gtgtacaaca caacaacaca
accaagaaag gttattctgg cattccattt attcagttta tttacagcc ttacttccag cagcacgtta aagatatggc cagggccggg tgcagtggct caagtctgta atcccaggac tttgggaggc caaggtggc ggatcacaag gtcaggagtt tgagaatctg gcaattcttc
                                                                                            72540
                                                                                            72600
                                                                                            72660
                                                                                            72720
agacttagaa gcaaccagct cgataacaca gtcttgtgtg ggctctccct ctgtccctcc
                                                                                            72780
cteqettece teatttetea tecetgeece tgagaetgtg caeetteaca tagecetgee
                                                                                            72840
atgagacett cateteagge tttgetttet ggggtaactg aggetaaaca etgagtggee
ctaaaagagg attgggattt ggaagttaga ttattcacca gagaacagac tttgctgatg
                                                                                            72900
                                                                                            72960
atcaggccca ggttgtaatt gttgaaaaaa agagaggatg catagtctta tctcatctcc
                                                                                            73020
tagtcaaagt caacaccatg ataaataaga gtcaaatcct gagatgtgaa ttggggacat
                                                                                            73080
ttgagtggtt aaccetgaga agettgeace tteagaceee teaataceee tgeteeeeag
agaaggotgg acattgacot cagcacaggo aggagocotg caagatgoca tttgtoctac
                                                                                            73140
taaagatgga cocctccact ctgtttctag gtaaataacc aaagtcaagt ctccacacag
                                                                                            73200
                                                                                            73260
cctgagcaag aaagtcagag cctgctacag gagaaaatac cacactggcc aaaggattca
                                                                                            73320
ctagccctgg ccactgtgtg tgggaggaac cagggaatca tgtgtgggag tcaatgttga
agetgttgga etgggggtgg ggtggaatat aageetggee etggggagtt ttteeegttt gagggeettt acceacaact caagatecag tgetatagea ggagatecca gagetagtee
                                                                                            73380
                                                                                            73440
                                                                                            73500
taacagatgg tcaggattga acttggccta gagtaaaatg aggaggatag tgccagaact
```

WO 01/27857

```
ttctcaacat actattgagg aagaggtcag aaggcttaag gaggtagtgt aactggaaag
                                                                                  73560
gggtcctgat ccagaccca ggagagggtt cttggacctt gcataagaaa gagttcgaga cgagtccacc cagtaaagtg aaagcaattt tattaaagaa gaaacagaaa aatggctact
                                                                                  73620
                                                                                  73680
ccatagagca gcgacatggg ctgcttaact gagtgttctt atgattattt cttgattcta
                                                                                  73740
tgctaaacaa agggtggatt atttgtgagg tttccaggaa aggggcaggg atttccaga actgatggat cccccactt ttagaccata tagagtaact tcctgacgtt gccatggcgt
                                                                                  73800
                                                                                  73860
ttgtaaactg tcatggccct ggagggaatg tcttttagca tgttaatgta ttataatgtg
                                                                                  73920
                                                                                  73980
tataatgage agtgaggaeg gecagaggte gettteatea ceatettggt tttggtgggt
tttggccggc ttctttatca catcetgttt tatgagcagg gtctttatga cctataactt
                                                                                  74040
ctcctgccga cctcctatct cctcctgtga ctaagaatgc agcctagcag gtctcagcct
                                                                                  74100
cattttacca tggagtcgct ctgattccaa tgcctctgac agcaggaatg ttggaattga
                                                                                  74160
                                                                                  74220
attactatgc aagacetgag aagecattgg aggacacage etteattagg acaetggeat
                                                                                  74280
ctgtgacagg ctgggtggtg gtaattgtct gttggccagt gtggactgtg ggagatgcta
ctactgtaag atatgacaag gtttctcttc aaacaggctg atccgcttct tattctctaa
                                                                                  74340
ttocaagtac caccoccogo otttottoto ottttoctto tttotgattt tactacatgo
                                                                                  74400
ccaggcatgc tacggcccca gctcacattc ctttccttat ttaaaaatgg actggggctg
                                                                                  74460
ggcgcggtgg ctcatgcctg taatcccagc actttgggag gccgaggcgg gcggatcatg
aggtcaggag atcgagacca tcctggctaa cacggtgaaa ccccgtctct actaaaaatg
                                                                                  74520
                                                                                  74580
caaaaacatt agccaggcgt ggttgcaggt gcctgcagtc ccagcggctc aggaggctga
                                                                                  74640
ggcaggagaa tggcgtgaac ctgggaggtg gaggttgcaa tgagccgaga ttgtgcact gcactccagc ctgggtgaca gagcgagact ccgtctcaaa aaaaaaaaa aaaaaaaaa
                                                                                  74700
                                                                                  74760
tagetgggea tggtggegeg tgeetgtaat accagetaet etggaggetg aggeaagaga
                                                                                  74820
atogottgaa cocagtaggo ggaagttgoa gtgagoogag atottgacac tgcactocag
                                                                                  74880
cctggtgaca gagtgagact ctgtctcaaa aaaaaaaaa agaaaaaaa agacagaaag
                                                                                  74940
aaagagcaca gacagagtca caggtatttg cagtaggaag ctgtcaggtt agagtgcacg gaaatagaaa gtatattta cacttacagc acatcttcgt ttgattagcc acatttaaaa
                                                                                  75000
                                                                                  75060
                                                                                  75120
tactgaatag caacgtgtgg ctatttagta ttcactaaaa tcttggacag tgcaagtcta
aagaateett gateegteeg geatggtgge teaegeettt aateeeagea etttgggagg
                                                                                  75180
ccaaggtgga aggatcactt aaggtcagga gttcgagacc agcctggcca acatggtgaa
                                                                                  75240
                                                                                  75300
acctcgtctc tactaataat acaaaaaaaa ttagccgggc atggtggtgc atgcctgtaa
                                                                                  75360
teceaggtae ttgggagget gaggeaggag aatagettga atecaggagg egetgeagtg
agccgagate atgccatgce actactgcae tecageetgg geaacagagt gagaetgtet caaaaaaaaa aaaaaaattg ttgggegtgg tggeteaege etgtaateee ageaetttgg gaggetgagg ggggtggate acetgggtte tggagttega gaccageetg gecaacatgg
                                                                                  75420
                                                                                  75480
                                                                                  75540
                                                                                  75600
tgaaacccca tctctactaa aaatacaaaa attagctggg cgtggtggtg ggcacctgaa
                                                                                  75660
atctcageta ctcaggagge tgaggcagga gaatttettg aacccaggag gcagaggttg
cagtgagcca agategegee tetgeactee atcetgggtg geagageaag actatgtete
                                                                                  75720
                                                                                  75780
aaaaaaaaa aaaaaaatac ttgattgtct ggacattctg cagaacatca tatggagaca
ctatgttgac gacatcatgc tgattgtaag caagaaatgg caagtgttcc agaaacacag
                                                                                  75840
                                                                                  75900
tcaagacaca tacatgccag aaggtgagat ataaactcta ctaagattca gtggcctgcc
acactggtga catttitaaa cotgotagat gtttgtgtag aaaaggattt aaccttgccc
                                                                                  75960
                                                                                  76020
aaagagggt ctggcctttg tccccagcta ctggacataa tctctttaaa ctcttgaaat
                                                                                  76080
atcattcctg atagaagtat ttttgttttg actaggggcc ttgggccagc cagatagcaa
caatgtgate tgggttgggg getttggate aggtggcate agtgtgacet cetgagtgge
                                                                                  76140
tagagactag aatcaaccac atgggcagac aacccagctt acatgatgga attccaataa
                                                                                  76200
                                                                                  76260
agactttgga cacaaggget tgggtaaget tteetggttg geaatgetet ataetgggaa
acceattetg actecatagg gagaggacaa etggatatte teatttggta cetecetggg etttgeecta tgeatttte eettgtetga ttattattat tattatgaga tggaateteg
                                                                                  76320
                                                                                  76380
ctctgtcacc caggetggag tgcagtggaa tgatctcaac tcactgcaac ctctgcctcc ccggttcaag cgattttcct gtctcggcct cccgagtagc tgggactaca gatgcatacc accacacccg gctaatttt ttgtatttt agtagagacg gggtttcacg ttagccagga
                                                                                  76440
                                                                                  76500
                                                                                  76560
                                                                                  76620
tggtctcgat ctcctgacct catgttccgc ctgcctcggc ctctcaaagt gctaggaata
catgtgtgag ccaccgcgcc cagccccctt ggctgattat taaagtgtat ccttgagctg
                                                                                  76680
tagtaaatta taaccgtgaa tataacagct tttagtgagt tttgtgagca cttctagcaa
                                                                                  76740
attatcaaac ctaaggatag ccttggggac ccctgaactt gcagttggtg tcagaaataa
                                                                                  76800
                                                                                  76860
gggtgctcat gtgtgtacca tgccctctaa ttttgtagtt aattaacttt cacaacttta
                                                                                  76920
ttattacege ttacaeteaa tgtttattea catttateea cataceaett attetagtge
cttgcatcaa agactttcta tctcatgtac tttattctgc ttgaagtaaa tcctttagga
                                                                                  76980
tattetttt tittttaaa etttqeacat acatactttt atttttatt tattttaat
                                                                                  77040
                                                                                  77100
tttgttattt ttgtgggtac gtagtagata tatgtattta tggagtacat gagatgtttt
```

```
77160
gatacaggca tgcaatgtga aataagcaca tcatggagaa tggggtatcc atcctctcaa
gcaatttatc cttcaagtta caaacaatcc aattacactc tttaagttat tttaaaatgt
                                                                               77220
                                                                               77280
acatttaatt ttgtattgac tagagtcact ctgttgtgct atcaaatata atttttttt
tttttgagac agagtctcac tcagtggccc agactgaaag tgcagtggca caagctcggc
                                                                               77340
tcacttcaat ctetgeetee etggttcaag cgaateteet geeteageet eccacatage
                                                                               77400
                                                                               77460
tgggattaca ggcacacacc accatgccca gctaattttt atatttttt agtagagacg
                                                                               77520
ggttttegee atgttggeea ggetggtett gaacteetgg eeteaaatga tetgaecaee
tcagcctccc aaagtgctag gattacaggc atgagccacc acacctggcc aaaatagaat
                                                                               77580
attetttagt gaggtetget ggtgacaatt tttttettt ttttgagaet gagteteget
                                                                               77640
gttgtcagct tgggctggag tgcaatagca cgatctcagc tcactgcaac ctccacctcc
                                                                               77700
                                                                               77760
cggattccag caattctcct gcctcagcct cccaagtagc tgagagatta caggcaccca
ccaccacacg cggctaattt ttgtattttt agtagaaatg ggggttcacc gtgttggcca ggctggtctc gaactcctga cctcaggtga tccacccacc ttggcctccc aaagtgctgg
                                                                               77820
                                                                               77880
77940
                                                                               78000
                                                                               78060
tcattattat tattattttg agacagggtc tcactctgtt gcctatgctg gggtgtagta
atgtgatete ggtteactge agaettgace tectaggget caggtgatet teccacetea
                                                                               78120
                                                                               78180
geeteectag tagetgggae tacagatgea tgeeaceata eccaactaat ttttetattt
tttgtagaga tgaggetttg ccacatttee caggetggte tetaaeteet gagetetage
                                                                                78240
aatccaccca ccttggcctt acaaagtgct gggccatgac tagccagcag ttacttttta
                                                                                78300
tagcatattg aatatttaat atgaatcttc tggcatccac tgtaactgtt taaaaaaatca
                                                                                78360
gctgtttact tggcactett ttttttttt tttttttga gacagagtet tgccctgtcg
                                                                                78420
cccaggetgg agtgcagtgg cgtgatettg getcaetgca agetetgeet eccgggttca
                                                                                78480
                                                                                78540
cgccattctc ctgcctcagc ctccggagta gctgggacta aaggcgcccg ccaccacgcc
                                                                                78600
cggctgattt ttttgtattt ttcgtagagt tggggtttca ccgtgttagc caggatggtc
tegatetect gacetegtga tetgteegee teggeetece aaagtgetgg gattatagge gtgagecace gegeecagee tettttttt ttttttttag aeggagtett aetetgteat
                                                                                78660
                                                                                78720
ctaggotggt gtacagtggo gtgatotoag otoagtgoaa cotocacoto otgootoago
                                                                                78780
ctgccaaata gctgggatta caggtgcgta ccatcacgcc cggctaattt ttgtattttc
                                                                                78840
agtagagatg gggtttcacc atgttagaca ggctggtctc gaactcctgg cctcaagtga
tctgcctgcc ccagcctccc aaagattaca ggcatgagcc accgcacccg gccaagtagc
                                                                                78900
                                                                                78960
actectttga aggtaatetg etteceetae eeetageaat tittaaeaat tittetteat tittatitee tgaagtiitg tiattaataa tetgigigea gattiettig tatteetit
                                                                                79020
                                                                                79080
gtttgcagtt catagtgatt cttgaattag tgtgttggtt tctgttatca ccacaggaaa
                                                                                79140
                                                                                79200
attgtcagcc gttagctttt caaatatttc cttgctaaat tctctcttct cccctttcgg
tacaattgat ttgattaaaa ctaaaaccag ggccgggtgc agtgactcat gcctgtaatc
ccaacacttt gagaggctga ggcaggtgga tcacctaagc tcaggagttc aagaccagcc
                                                                                79260
                                                                                79320
tggccaatat ggtgaaaccc cgtctctact aaaaatacaa aaattaccag gcatggtggc
                                                                                79380
                                                                                79440
acacatttgt agtcaggagg ctgaggcagg agaattgctt gaatccagga ggtggaggtt
gcagtgaget gagateccae caetgeagte tggeetggge gacagagtga gatgagaate
                                                                                79500
tgtctcgaaa aaaaaagtta tgaatgtttg ataaactata tttgttagaa tgtttgttgt
                                                                                79560
agaatactat tcattgattt ttaaacaatg ttagattaaa ccattcactg gatttgtgat aattaactta ctgattttac ctcactgatt tgttgtaatt aatacaactg gtataaaaag
                                                                                79620
                                                                                79680
                                                                                79740
actgtgacga ggccgggcat ggtggctccc gcctataatc ccagcacttt gggaggctga
                                                                                79800
ggcaggcgga tcacctgagg tcaggagttc aagaccagcc tgaccaacat ggtgaaaccc
                                                                                79860
catctttact aaaaatacaa aattageegg tegtggtggt geatgeetgt aateecaget
cttcgggagg ctgtggcagg agaatcactt gaacccggga ggtggaggtt gcagtgagcc
gatatcgcgc cattgcactc cagcctgggc aacaagagcg aaactccgtc taaaaaaaaa
                                                                                79920
                                                                                79980
aaagaaaaaa aacacataaa acaaaacaac actgtgacgg ttcccaaaaa ttaggagcat
                                                                                80040
                                                                                80100
aattaaagga actcctgata aaaattaatt ttatcttaca tgtaaactaa aatgacttta
                                                                                80160
tgaagttaat tcagaaatac aatgcagggt attagtttgc cacagetgcg tattcagect
aatgtaatat tottgttatt tttaaattot tottttaact ttactcatat gtggatcatc aaatttcaaa agattaaatg acaatactot tagcagcaag ottooctaag catataaaca
                                                                                80220
                                                                                80280
                                                                                80340
ttttaatggg tgatgattca gaaggtaccc gaagaatatg tactgccaga tatcattcac
                                                                                80400
ccccatatac ctgcccgaca gacateccat tttgggaccc tggataaatg tgtgggtgga
gagaaagata ggagaaagtg gtataagcaa atggctttgg agtctgattg acagcgattg aaatcctgtc tctacctctt aacagcctca tgatcctaca taagttaccc cgatcctcag
                                                                                80460
                                                                                80520
ggccacatet gtaaattggg ggttgcgatg gcagccatet cacagggtet citttegggg
                                                                                80580
                                                                                80640
aagggcagga attatggatt aagtgagcta gtaattgtaa agcacttaat acaaggaggg
                                                                                80700
cgcataataa gtacttcata aataatgacg gccattatca tgactgaggt gtatgcagct
```

WO 01/27857

```
gtcggggatt acggcgactt cagaatttct ggtgggcagg gctcaaaggc agcaaatcac
                                                                             80760
actggaagtc gaggtgaggc actgettetg cacagactge ttagetggag agaatgagga
                                                                             80820
aggettagag gagatttaga ggaacttaga gteeteegee teeaactetg tgggatetge
                                                                             80880
                                                                             80940
tecegtgeca gagacattea ggggatttet egeactetee ceteceetae gteceteeeg
                                                                             81000
ccccatccaa ctaaccacac aacacataca aaatagcccc tgcgaggttc tgcacgctgg
aagggaacag gagaagggcg ctgcgctttc ttgctgatgc cctgtacttg ggcccctggt
                                                                             81060
agacacagee actigiecee teageeigea gagaaateee aegiagaeeg egeeegggie
                                                                             81120
cttggcttca gccaatctcc ctttggtggg ggtgggatgc acgatccaag gttttattgg
                                                                             81180
                                                                             81240
ctacagacag cggggtgtgg tccgccaaga acacagattg gctcccgagg gcatctcgga
tecetggtgg ggegegete ageeteeegg tgeaggeeeg geegaggeea ggaggaageg geeagaeege gteeattegg egeeagetea eteeggaegt eeggageete tgeeageget
                                                                             81300
                                                                             81360
getteegtee agtgegeetg gaegegetgt cettaactgg agaaaggett cacettgaaa tecaggette atceetagtt agegtgtgae ettgageagt tgaetttatt ttteagtgee
                                                                             81420
                                                                             81480
                                                                             81540
tagttttcca gataccagga ctgactccaa ggactattac tcatctggag ggtttagcac
agtaccgtcg catagtaaat ttccatgtca gttttggtta cctttcatgc acttgcaaac
                                                                             81600
atgccatgct ctgaaacgaa ataggcacat ctttttttt ttttttta aggagtcttc
                                                                             81660
ctctcgccca ggctggagtg cagtggcgcg atcttggctc actgcaacct ccacctcccg
                                                                             81720
                                                                              81780
tgttcgagat tctcctgcct cagcctcctg attagctggg actacaggca tgccacgacg
                                                                              81840
cccagttaat ttttgtattt ttagtagaga cggggtttcg ccatcttggc caggctggtc
                                                                              81900
taacteetga ceteaggtga tetgaetgee teageetete aaagtgttgg gattacagge
ataagccact gcatctggcc agaaatgaaa taagtaaatc ttttaacctg ctctaacaat
                                                                              81960
atagtgaaaa gaccatatta ttattagagc aggttaaggg atttgcctat ttcgggttct
                                                                              82020
                                                                              82080
agttatagtc ttaaacttgg acattettgt agaaagtaaa aagttteete tteaaagtte
                                                                              82140
cccttcttgt taaagaatac atcataagtg ttagaagtaa tagtttattt taaagactaa
ctttcttcaa gcctccttgc tttgtgctaa taactctttg ttaagcccta tcctatgtaa ctgttggaca tgctcacagg cacgttccag ttcacagcct atgccccttc cttatttgga
                                                                              82200
                                                                              82260
                                                                              82320
aatgttattg cttccttaaa cctttcggta agcaacttcc tctccttctt cgttcttcct
                                                                              82380
tgcacttacc tatttagaaa gttttaggct attagcaaat cggctatcag tttaagagtg
                                                                              82440
tgaggteceg etecagecaa tggatgeagg acatageagt gaggaegaee caaatgegta
agggataaat atgtttgett tteetttgtt caggtgtget etegacateg tteeatetge gattgageae cetttetgea gaaagtaaag attgeettge tggagatett ttgteteegt
                                                                              82500
                                                                              82560
                                                                              82620
gctgactttt cttcgtggca ccgattatct atttctaaca attttggtat ttctaacatt
                                                                              82680
ctgaacaatc ttgggctagt tgtctcttct gggcctgttt ccccatccgt cacatgataa
acttcattgg tttaaaaacc ccagcgaaca tttattgagt tactattacc ttcctgccct
                                                                              82740
ccccaaccc aaccccaggg agcagttaca acctcagccg ctgagcgcac tcgccgggtg
                                                                              82800
ttaagaagca ccaaagacag ggaggettga ttgattttgc tttgggagta gagggtcaga
                                                                              82860
agattcacag gaaaatggca tttgagcaag gatgattcac tggagctagc ttttaaatac tggcgaggct tttatgttgc agtcccttac aaagttgagc attcgcaggg actgcactcc
                                                                              82920
                                                                              82980
gaaataagcc cgcttcccct tttcattcgc taatgatcca gggagctgct ggttccgcat
                                                                              83040
geggeaggtt gtgeetttte etaateaggg ttetgeateg cetegaacee geaggeegtg
                                                                              83100
gegggttete etgaggaage agggaetggg gtgeagggtg aagetgeteg tgeeggeeag
                                                                              83160
                                                                              83220
cgcctgtgag caaaactcaa acggaggagc aggaggggtc gagctggagc gtggcagggt
                                                                              83280
tgaccetgee ttttagaagg geacaatttg aagggtacee aggggeegga ageeggggae
ctaaggcccg ccccgttcca gctgctggga gggctcccgc cccagggagt tagttttgca
                                                                              83340
gagaetgggt etgeageget ceaceggggg eeggegacag aegeeacaaa aeagetgeag
gaaeggtgge tegeteeagg cacceaggge eegggaaaga ggegegggta geaegegegg
                                                                              83400
                                                                              83460
                                                                              83520
gtcacgtggg cgatgcgggc gtgcgcccct gcacccgcgg gagggggatg gggaaaaggg
geggggeegg egettgaeet eeegtgaage etagegeggg gaaggaeegg aacteeggg gggeggettg ttgataatat geeggetgga getgeetggg eateeegagg aggeggtggg
                                                                              83580
                                                                              83640
                                                                              83700
gcccactccc ggaagaaggg tcccttttcg cgctagtgca gcggcccctc tggacccgga
agteegggee ggttgetgaa tgaggggage egggeeetee eegegeeagt eeeeegeae
                                                                              83760
                                                                              83820
cctccgtccc gacccgggcc ccgccatgtc cttcttccgg cggaaaggta gctgagggg
cgccggcggg gagtcaggcc gggcctcagg ggcggcggtg gggcaggtgg gcctgcgagg
                                                                              83880
                                                                              83940
tgagtgcctt gctctgctcc gggactcttc tgggagggag aaggtggcct tcttgcgcga
                                                                              84000
ggtcagagga gtattgtcgc gctggttcag aagcgattgc taaagcccat agaagttcct
                                                                              84060
gcctgtttgg traagaacag trottaggrg ggggttagtt trrttgtgtt totttgagga
                                                                              84120
                                                                              84180
ccgtggatca agatcaagga aatctcttta gaaccttatt atggaagtct gaagtttcca
aatgitgagg gitttatgic taaaagcaac acgtgaaaaa attgttitct icacccagtg
                                                                              84240
ctgtcttcca atttcctctt tggggggagg ggtagttact gctgttacta aaataaaatt
                                                                              84300
```

WO 01/27857

```
acttattgct aaagttcccc aacaggaaga ccactacttt tgatgacttt ggcaagtttg
                                                                                          84360
                                                                                          84420
ctaactactg gaaccctaac ttacaaacga actacttaca tttttgattt ccagttgtat
tacctgccca atgtttacgt agaaacagct taattttgat totgggtaac gttgttgcac
                                                                                          84480
ttcattaaaa atacatatcc gaagtgagca agtatgggtc tgtggacagc agtgatittt
                                                                                          84540
cctgtcaatt cctgttgctt cagataaaat gtaccagaca gaggceggge geggtggete acgectgtaa teccagcact ttgggagget tggegggtgg atcacetgag ategggagtt
                                                                                          84600
                                                                                          84660
                                                                                          84720
caagaccage etgaccaaca tggagaaace cegtgtetae taaaaataca aaattageca
                                                                                          84780
gggtggtggc gcatgcctgt aatgccagct acttgggagg ctgaagcagg agaatcgctt
gaacctggga ggcggaggtt gcggtgagcc gagatagcac cattgcactc cagcctgggc
                                                                                          84840
                                                                                          84900
aaaaagageg aaacteegte teaaaaaaaa agtaceagae agaaatgggt tttgttttet
                                                                                          84960
ttttttgttt tgagacggag tttcgctctt gttgcccagg ctcgagtgca atggcgcgat
                                                                                          85020
ctcagtctcg gctcactgca acctctgtct cccaggttta atcgattctc ctgcctcagc
ctcccaagta gctgggatta cccatgccc accatgcccg gctaattitt gtattittag tagaaacggg gcttcaccat gttaggctgg tcttgaaccc ctgacctcaa gtgggcctcc cacctcggcc tcccaaagtg ccaggattac aggcatgagc caccgcggcc agccagaaat
                                                                                          85080
                                                                                          85140
                                                                                          85200
gggttttgga aaaagcacta aacaaaatcg aacttggttt catatgacag ctctgctgct
                                                                                          85260
aactgtaaca ggggcagacc agttaaccta cttttctgtc ttctgtcagc tgagaattag
                                                                                          85320
atgattecca aaggeecatt gaactetgaa tgaetttaaa taettettet taagtgggta
                                                                                          85380
cacggttttg gtaactgatg ccaggtgatg aatgcatgaa agtgcttaat gaatgaaacc ggtaaaatag taggaggaag ctttattggt aaggcagggg tatacctaat agctctctaa
                                                                                          85440
                                                                                          85500
tttattggta ttgaagtggt taacttttgt ttttttaagg ggggaaaaca ttctaagaat aatgaggcaa actgcatatt gcacaagaga ctgttgtctc tattcaacaa ataccttttg
                                                                                          85560
                                                                                          85620
agtgtccaga gtctgccagg tgctgtgcta ggccctcacg attgagtagt gaaccagaga
                                                                                          85680
atgtccctgc acccatggag cttattgtct actggggtag acagataata aataagcaaa
                                                                                          85740
                                                                                          85800
caaatcttct ctcttctccc tttcgctcca tgtaagtgtg tgtgtatagg tgtatactta
caagttgagt aaagtgttat gaaagattaa gaggagaaat gcattttggt tagatgttag
                                                                                          85860
aggactcage aggtgacett gaaacttaga getgaaggat cagtaggagg taactagaga
                                                                                          85920
ggccagggaa tcgcatgttc aaaggccagg aggcaagaaa gagcatggtg cccttcaaga gaggaaagaa ggctactgtg actggagcat agatgtaggc aagtgttggg tgattgagag ctctacgggc catggttagg ttttattcct aatgccgaga tgccaaacat ggtggttcat
                                                                                          85980
                                                                                          86040
                                                                                          86100
atctgtaatc ccagtatttt aggaggccga ggcaggaata tagcttgaac ccaggagttc
                                                                                          86160
aagaccagcc tgagcaacat gagacctgta caaaacattt aaaaaattgc tgggtatgat
                                                                                          86220
                                                                                           86280
ggtgcacacc tgtggtccca gctactcagg aggctgaggc agaaggatca cttgagccta
ggaggtggag gctacaatga gccatatttg agtcactaca ctccagcctg gatgacaaag tgagaccatg tgtcaaacaa aatacagaaa gaatattaat ttaaaatttt gaaagaggag
                                                                                           86340
                                                                                           86400
tgatetgaae ttatatetta aaaagateat tetagggeat ggtggeteat geetgtaate
                                                                                           86460
aagggetttg ggaggetgag acaggaggat cacetgagge cagttegaga teaacetgta
                                                                                           86520
                                                                                           86580
cagcatagag agactccatc tctacaaaaa gaaaaaataa atagctgggt gttgtgagtt
attcaggagg ctgaagcaga aagatcactt gagcccagga gtttgaggct gcagtaagct atgatcccac cactgcaaca cagtgagatc ttgtctcaaa aaaaaaaaa aatcattcta
                                                                                           86640
                                                                                           86700
                                                                                           86760
ggtgcttttt ggaggctgga tgtggtaaga gtagaagctg gagatggtcc tgttagggat
tcgattcaga ctttaaatac catcaatgca ttgagtccca aatttacatc actacgttgg
                                                                                           86820
                                                                                           86880
atcettgeee etgaateeag actggtatat ecaaetttag gtteagtttg tatetetace
tgaccaatat agaggtgtcc agtcttttgg cttccctagg ccacattgga agaagaattg tcttgagcca cacatagagt acactaacgc taacaatagc agatgagcta aaaaaaaatc
                                                                                           86940
                                                                                           87000
gcaaaactta taatgtttta agaaagttta cgaatttgtg ttgggcacat tcagagccat cctgggccgc gggatggaca agcttaatcc agtagatacc ttcaacttac aatatctaaa
                                                                                           87060
                                                                                           87120
                                                                                           87180
attitiatgcc agatttagtc attttaaacc tgctcatcag tttttctcaa gaagtagtat
                                                                                           87240
tttggctitt titctttict tttttttgag alggagttte getettateg itcaagetgg
agtgcagtgg cggatettgg etcaetgcaa ecteegeete etgggttcaa gtgattetee
                                                                                           87300
tgecteagee tegeaagtag etggaattae aggeatgege caccatgace agetaatttt
                                                                                           87360
tggagacagg gtttcaccat gttggtcagg ctggttttgt actcctgacc tcaggtgatc
tgcctgcctc ggcctcccaa aggctgggat tacaggcatg agccaccgct cccggctgca
                                                                                           87420
                                                                                           87480
                                                                                           87540
tttttggatt tttagttgct cagcccaaaa ctttagtaca tctttgaacc tcttctttcc
tectacteta tatetgatee ateageaaat etgttaggte taceteacae atategaaat
                                                                                           87600
                                                                                           87660
cctaccacgt ctcaccatct gtgacaatta acaccctggt ctaggcagtc atctctgtta
agattgagtg gttaaggatg teetetaagg agatgaeatt caaatettag ettaaatgte aagaggage tggttttata aagattgagg aggeageatt attttgeeat aggetteeat
                                                                                           87720
                                                                                           87780
ttggtttcca ttccattctt gatacttatg gtatatattc aaaacaaatg cacagaaaca gacccaggta tattgggaat ttcggatata gagttcctag ttgggaaaag atagactgat
                                                                                           87840
                                                                                           87900
```

WO 01/27857

```
ctgtaaatga tgctagttat ccatcatctg gcaaaaaata atttcctgcc tcctctcata
                                                                                                              87960
tateteagat caacagaett tttetgttaa gggccaaate ataaatattt taggetttee
                                                                                                              88020
agaccatatg gtttctgtca cactctcctt tatccttgaa gccatagaca atatgtaaac
                                                                                                              88080
                                                                                                              88140
aaatgggcat ggctgtgcta cgataaaact ttacttacaa aaactggtag tgggccagtt
taggcatgge cagcactttg ggaggctaag gcagatggat cacttggggt caggagtttg agaccagect ggccaacatg gtgaaaccet gtetetacta aaaatacaaa aaatagetgg gcatggtggt gggtgtetat aattecaget actetggagg ctaagacaca agaatcactt
                                                                                                              88200
                                                                                                              88260
                                                                                                              88320
gaacccagga ggcagaggtt gcagtgagct gagatagcac cactgcactc cagccagggt
gacggagtct taaagcaaaa caaaacaaaa ggtagtgggt tgtatttggc ccatgggctg
                                                                                                              88380
                                                                                                              88440
tagtttgcca atecetgatg cagaaacaaa ttecaggtaa ataagageet ggaatgttaa
                                                                                                              88500
aaaaacaaaa cttgaagtca tgtagaagaa caggtagggg gaacaatcct gatctcagga taggaaggga tattgcttaa aataagacac aggaaaatat aatccatgtt gtgtaaattt
                                                                                                              88560
                                                                                                              88620
                                                                                                              88680
gactacgita aaacttaaaa ctttcgccaa gcgcggtggc tcacgcctgt aataccagta
ctttgggagg ccgaggtgag cagatcacca ggtcaggaga ttgagaccat cctggctaac
                                                                                                              88740
acggtgaaac cccgtctcta ctaaaaatac aaaacattag ccgggcgtgg tggcgggcgc
                                                                                                              88800
ctgtagtccc agctacttgg gaggctgagg caggagaatg gcctgaaccc gggaggcgaa
gcttgcagtg agctgagatc gcgccactgc actccagcct gggcgacaga gtgagattcc
                                                                                                              88860
                                                                                                              88920
gtotcaaaaa aacaaaacaa aacaaagcaa aaaacctaaa actttcatac aataaagtat
                                                                                                              88980
                                                                                                              89040
acctaagata cttctagaag agaagattta catccaggac gtgtatggaa tttctgcaag
taataagtaa aagacaaggg acatgaagag gcagttcaca aaagaggaag ccaaaatgac caataaacat gaaaggatgt ttaacctcaa aggaaacaag gaaatgaatt aaaaacatca aatgccattt caaaactagt aagttggcaa aattaaaaat accaaggatg agaatatgaa
                                                                                                              89100
                                                                                                              89160
                                                                                                              89220
gcatggctat atgagtgcat ggaatggtac agtcactttc attaaaaatg cacataattt
                                                                                                              89280
gettettatt tattetttg agacagteta tgtegeceag getagaatge agtggeatga teteggetea ccacaatete tgeeteetgg gtteaageaa tteteetgee teageeteet gagtagetgg gattacagge acatgceaca acgeeeggtt aagtettgta ttettagtag agacagggtt ttgecatgtt ggeeaggetg gtetegaact cetgacetea ggtgagetge
                                                                                                              89340
                                                                                                              89400
                                                                                                              89460
                                                                                                              89520
                                                                                                              89580
ttcccaaagt gctgggatta gaggcgtgag ccaatgctcc tggctgaaaa aaatgcacat
aatttgttac ctagcaattc catgtctaga ggcttatcct agagaaattc ttgcttatat gcataggaag acgtgtacta gaatgttcac tagttgaatg tttaagtgaa aattaggaaa
                                                                                                              89640
                                                                                                              89700
taaagtaaat gttcattaac aggaaaatga gtaaaggtat atttataaaa caattaagta gctaaaatga ataaactaga gctgcgtgaa tgaactagaa ctggttcaat agtcatgtca
                                                                                                              89760
                                                                                                              89820
gattattgaa tgaatacagg tcagatatgt atagagtgtc attigtgtaa ttaattittt
                                                                                                              89880
ttttttttt gagatggagt ctcactctgt tgcccaggct ggagtgcagt ggcgtgatct cagctcactg caacctccac ctcctgggtt aaagtgattc tcctgcctca gcctcccgag
                                                                                                              89940
                                                                                                              90000
tagttgggat tacaggcatg caccaccatg cccagctcat tttcctattt ttagtggcca cagggtttca ccatgttggc caggctggtc ttgaactcct gacctcaagt gttccaccca
                                                                                                              90060
                                                                                                              90120
acttggcctc ccaaagtgct aggattacag gcgtgagcca ccgtgctcag ccatttgcgt gatttttaaa gatgtgcaga ataatgccat taaaaaaaat acacatacat gtatatatat
                                                                                                              90180
                                                                                                              90240
acacgtttgg ctgggtgtgg tggctcacac ctgtaatccc agcactttgg gaggctgagg
                                                                                                              90300
                                                                                                              90360
caggaggate acttgagece aggtgtacaa gactageetg ggegagatag caagaceeca
totcaacaac agaaaggata attaggtatg gtggcatgag aggatcactt gagcccagga gttcgagtgt tatcaggcca ctgcactcta gcctggacaa caaagcaaga ccgtgtctca
                                                                                                              90420
                                                                                                              90480
aaaaaataaa aataaaaagt atttgtatgt ggtcatagtc aaaaaacgta catggaagga aaatgtcttt atttatttat ttatttttt ttttttaaga cagagtcttg ctctgtcacc
                                                                                                              90540
                                                                                                              90600
caggetgggg tacagtggtg taateteage teacegeaat eteggeetee egggtteaag egattettet geeteageet tetaagtage tgggaetaea ggtaceegee accacacet getaattett gtgtttteag tagagaeagg gttteaceat gttggeaagg etggtetega acteetgace ttaagtgage caceegeett ggeeteecaa agteetggga ttacaggtgt gageaactge gettggeeag gaaatateta atttagtaag tatttatate tgggaaagga
                                                                                                              90660
                                                                                                              90720
                                                                                                              90780
                                                                                                              90840
                                                                                                              90900
                                                                                                              90960
agggtcaggt ggtgattcat aggaacteta aagtetatgt ataatactta gggggacaga
aggaaataaa gcaaaatgct gatatttgat tgttgagttg tgtatatgtt agaagtataa
                                                                                                              91020
cataggagat ctgattgata gtaggagaat gtttttaggt ggtaaaagtg gaaccgtggt
                                                                                                              91080
ggtttgtttt ggcagtagaa tcagttggtc atagtttgta tgtggaaggt aataaacaga
                                                                                                              91140
                                                                                                              91200
ccatgttaag gatgacticc ggaattitigg totgagtagt gggtggatga cagtgtcatt
catgagggaa gatgaagact gaggtaggaa caggtttggg agaagatgac atgttccctt
                                                                                                              91260
ttagacaagt ggaattatgg aagatggcag gtaggtggtt agctatatga atttgagata aaagatttag gatggagata taaatttagg agtaacagcg tatctatggt attgtaagcc ttaagaatgg gtaggatcag ccaggaaata cagatgtata tgcagaagag aggagtcaag gaagccaaga caagttaatg tttaaagtga gtgatgtagt ccatgggcag atgctgctga
                                                                                                              91320
                                                                                                              91380
                                                                                                              91440
                                                                                                              91500
```

```
gagggetgea aacaccagtg accetacaac atttttaaat gtegtettee tgacagcagt
                                                                                                    91560
gatcagtacc tgcaacgatc ttatttattt ttttcatgtt agtctccaca cacttgaatg
                                                                                                    91620
tagacttttt gaaggcaaaa tcattgcctt ttctgagctg ggagcatgtc tggcacatac caagcactca acagttgatg tattgacttc atccagatac tctgagggcg agttatttcc
                                                                                                    91680
                                                                                                    91740
                                                                                                    91800
tgctactage etticacett teaatgttta agageacaaa tacagagatg ggcacgtttt
ggcatttctt attttgataa ccttttcctg gtaagatttt ttaatgttga aaaaaaaaa caagaaaaaga gggttaaaaa tagtcttatg tcagatcctg tgatagaatt cacacttggc
                                                                                                    91860
                                                                                                    91920
ttaagetget gggeacette etatettgga tgteatatta gettatetae ageagaattt ttaetgtttt atgtagtaag gaageaatta tatgattatt ttaeagaeaa attattettt
                                                                                                    91980
                                                                                                    92040
atcttttatt tttttagacg gagtctctct ttgtctccca ggctggagta cagtgtcgcg atctcggctc actgcaacct ccgcctcctg ggttcaagca attctctgcc tcagcctccc aagtagctgg gcttacaggt gtccgccacc acacccagct cattgttttg tatttttagt
                                                                                                    92100
                                                                                                    92160
                                                                                                    92220
agagatgggg tttcaccatg ttggccaggc tggtcttgag ctactgacct caggtgatcc accepecttg gcatcccaaa gtgctggaat tacaggcgtg agccaccgtg cctggcccag acaaattatt atactctgag tgttagaggc ttaggatgtt ttcacttgat gctatgggag
                                                                                                    92280
                                                                                                    92340
                                                                                                    92400
gaataagtaa taagatatga tacacaacca aagacettte tteactatge ttetagtage
                                                                                                    92460
tagtactatg gatgacacat ggtaataata ttggttagca tttgtcctca atttactgtg
                                                                                                    92520
ctagttacto ttotaagooc ottacaggta tatattttt ttoatcaata atcototaag
                                                                                                    92580
gtagttttta ttattgacct aattttataa atcaagaaaa ttaagaccca gagaagtaag
                                                                                                    92640
taacttgtce aagatcacat ggettataag tggtagagec agaatttgac eccagatgtt gtgactacat tgtetetea taageaggtt caactetttt gaetggatge tgttecaagg teaetteett agagaageet ttgetgacaa etaeceteet gtgeeeteet ecaaggetgt
                                                                                                    92700
                                                                                                    92760
                                                                                                    92820
                                                                                                    92880
ccattgttct agaactttga atactcatct tagaataaag ctggtctaat ttttacagtg
ttatagaatg gatctctgac tgcaaaagtt ggtcataatt atctttttat gttctagtga aaggcaaaga acaaggaaag acctcagatg tgaagtccat taaaggtaag ttctgcctt ggcagtccac tgcattaaaa agtgatgtgc tttgcatttg tgagttctt aatcctgtta tactctctct tttggcatta atcatttctg ccttatttta taattactta tgattttgat
                                                                                                    92940
                                                                                                    93000
                                                                                                    93060
                                                                                                    93120
ttatttccct ctttaacctg tataatgctt taacatctag catataataa gtaggctttt
                                                                                                    93180
ttttttttt ttttttgga gacggagtet tgetetgtta cecaggetgg agtgeagtgg egegatettg geteactgea agetetgtet eeegggttea caccattete etgeeteage
                                                                                                    93240
                                                                                                    93300
ctccccagca gctgggacta caggtgcacg gcgccacgcc tggctaattt tttgtatttt ttagtagaga cagagtttca ccatgttagc cagtatggtc tcgatctcct gaccttgtga
                                                                                                    93360
                                                                                                    93420
teegeegee teggeeteee aaagtgetgg gattacaage gtgageeace geaceeggee gtaagtagge ttttttace ttaattttat ttttttgaga tggagtettg etettateee
                                                                                                    93480
                                                                                                    93540
caggetggag tgcagtggtg ccatctcggc tcactgcagc atccacctcc cgggttcaag
                                                                                                    93600
                                                                                                    93660
cgatteteet geeteageet eeegagtage tgggattaca ggtggeegee accatgeeea
                                                                                                    93720
gctaattttt gtatttttag tagagacagg gtttcaccgt gttggccagg ccagtctcaa
actectgace teaagtgate cactegeett ggeeteecaa agteetggga ttacaggegt
                                                                                                    93780
gagccaccat gcctggccat aagtaggctt ttactgagcc ttgtgtgtat tggctatcct
                                                                                                    93840
agtgattaca gtgaaccagt gcccttctta ttaatcacac atttaattgt tccctaaaag tgattagttc actttattta tttagtaaga caaaaaatga agaatactct taactgagca gtctgttaac tgtaggaaag cactgacact tataaggctt agtttctgt catttatcca
                                                                                                    93900
                                                                                                    93960
                                                                                                    94020
gaagtatggt tgattacagt ttttactttt ttatttgaat gaacaacctt aatttaaaat
                                                                                                    94080
atattttgtt tatttttgt tgggatcgat acattgtcct tgtttataga ttagagcatg ctttttaaag atgctgtatt actcactgat tttatttgtc cagtgtacag agattgaagt
                                                                                                    94140
                                                                                                    94200
gggaaaatta taatggaaat tgtttccata gtcattacat attaatttca tcaatttatt
                                                                                                    94260
tccataaaat ctgtagattg ctacttattt agatttttcc ttcaaatgtt tttatgttgt
                                                                                                    94320
attgcttgca ctgagtattt attctatatg ctcaatttgc tggagaagaa gactaattat
                                                                                                    94380
aacttaggca agttgtaaaa ttagggaaaa aagtaaggta ccttacagcc tagtttactt atttcttatg taaagccagt tagattccac attagttcaa actgccttct ttgagcaaaa
                                                                                                    94440
                                                                                                    94500
                                                                                                    94560
cttgattggc agtgataaag gcttaaagcc cttctcaagc agagacctgt aaagactaga
totgactgta gtagaaggaa ggaacttaga tgtttcaggc agtgagaaca ccagtottco
                                                                                                    94620
actotaaact ttgccactaa cagtatgacc ttgggaagtt gtaactttct tcagattott
                                                                                                    94680
catttgttga atggggggat tggcctagct aatttctaaa tctctactgg gctaaaaaat
tctgtgctta tactctgatt atgaagtaca taatctgtgc ttaacattca ctgacttatc
                                                                                                    94740
                                                                                                    94800
cttaggataa tacagaagca gtacaagaaa cagcccttca agatgtttgc agtctggtta
                                                                                                    94860
gaaagacaaa cttatacaca gaacagtagc aaatagacca aaataataat agctgccatt
                                                                                                    94920
tatagaacac ttettetgtt etgggeatta gacaaaaact gactataacg gtgaacaaaa
                                                                                                    94980
aagacttagg teetgeeete attgaaetta cagattagta ggggagagga acattaatca
                                                                                                    95040
agtaattcca cagatggctt agcctagatt ggtagtgatg gaagtaaaga gatgtgaacg
                                                                                                    95100
```

```
gacttgaaaa aaaattcgga ggcaaaatgg atagaagttt attattgatt aaatatgagg
                                                                         951
tgtgagagag agggatattt aagattgata cctaccttct ggcttgccta acagaaccaa
                                                                         952
aacaggaaat tatatgttca gttttgttat gttgggtggg aggtgctttt gagtcattca
                                                                         952
tttatatatg ttatatatgt tattttatat gcatagtaat tttaaggtct gagttttaaa
                                                                         953
                                                                         954
ccaaaggtta gagagtgatt ttttagagtc tagcaaacct aagttgaaat cctgcctgtt
gaaatggctg tttactagct cattaaccta gggcaaagta ttcaacttgt tttcattttt
                                                                         954
gtcttcatct ctaaaatgag gaaaatatgg tcttacaaga ttgtcctgag agatagatga
                                                                         955
aataatatee aaaaaaaaa aaggtacata gagaaaeteg tatagtgeet ggtatatagt
                                                                         955
aggtoctoca ttggtagota toattatota gtittaacai agcottoagt tigttgaatt
                                                                         956
agtcaaactg agtgaagcac tgcaaggaat tcagaggaat ttgagatcaa caaatgattt
                                                                         957
ctgaagttta gggaagactt catggcaatg acacttacct tgtataaaag ttgaagaata
                                                                         957
agaaagattt gaatgagaga ttetttetet tetecetace ageceagett ettatttgag
                                                                         958
gatatattgg gcaaaggggc cttcagacaa gtagagggag atttttacag aaagattgag
                                                                          958
atgaaggtat agaaggctgt aaagaccaga aaagagaatt gagacagagg aagcaggaag
                                                                         959
ccactgtagg titttgagca agatattgat gctgtaagta tggtgtttat gaaaggttag
                                                                         960
tetggaagag atttgcagga tggagacece ggaagttttt ttgttataat acagaaagac
                                                                         960
ttgcactgag ggtgaggtgt taaaaataaa caggtaagta aatgtttaaa catcttgaag
                                                                          961
gaaaagtcaa caaatcttgg caagtaaaca gataacagtg aaaaagaatg ggaccaagat
                                                                         961
tttgagtttt ggagactggt ggattgaaca gacagggaaa ttgagaggag aatcagatga
tgatgttta agttgatatt tagacagatt gtgcttgaga tggtaaagtc aatgtgggtg
                                                                          962
                                                                          963
ggaatgetta gtagegagta atcagtgata caagaccaaa geecaggtea aagacaagte
                                                                          963
                                                                          964
acagatacag atcagggett tttcatetge tecacagagg tgtaccetag gagetgttge
aaacagteca tgtggagggt gtgagtaaga tgttteeett gaatttgeea gaattaettt
                                                                          964
tttgttgttg ttgttgtttt ttctgagaca gatteteget etgttgeeca ggetggaggg
                                                                          965
cagtggcgag atcgcgcagc tcactgcaac ctctgcctct cgggttcgag tgattctcct
                                                                          966
geeteageet eccaagtage tgggattaca ggettgtgee accaageeea getaatttet
                                                                          966
tttgtatttt tagtagagat ggggtttcac catgttggcc agactggtct cgaactcctg
                                                                          967
geetegtgat etgeetgeet cageetecaa aagttetggg attacaggeg tgaaccaetg
                                                                          967
 caccoggico etigitaagi tiattitggi gggaagcaaa ggaggittca gcttitaaaa
                                                                          968
agtttgaaaa ttattgetet ggtaataatt aaagatttga gagtaaatat getttetage
                                                                          969
                                                                          969
 agaaagaata aaagaagaac agatageete aagaagggga gecaaagaag caggetatat
 ctgacacact gggtgttgat aaatgggtat taaaagaatg agagcaatga gcagatagaa
                                                                          970
gaggaaatta ggagagtata ataccatgga gaccaagaaa gatagactat caggaaggag
                                                                          970
 tggtaaaaat aagttactag ttctaagaga gatgttaaga gggaccgggg aaagccttgt
                                                                          971
                                                                          972
 acaaatgagt tagtagcatt ttacattata tacatctaat taagaaacaa
                                                            tgcgagagtc
 tcaccattcc tatagactct tacttgtact tgtctgaaca cgaaaactgg
                                                                          972
                                                            cttttgttta
 taaataaget aaaaattatt ttgeteeaat tteteatgaa aataaaaata aacettettt
                                                                          973
 taacattgaa aaaatagttt gaagacagtc actcttcatt ttgtaattcc cacaactatt
                                                                          973
                                                                          974
 attgaatgac tgaaattatc tttattctga agccaaaggg gtgatactga tatttcttca
 gactactaaa aatatattt atgaatttt agtgtgcttt atctttttt gtttttttt
                                                                          975
                                                                          975
 ttgagatgga gtttcactcc cgttgctcag gctggagggc agtggtgcaa tctcagctca
 ctgcaacett egecteceag attcaageaa tteteetgee teggteteee aagtagetgg gattacagge acetgeeee acacecaget aattttttgt atttttagta gagacagggt
                                                                          97€
                                                                          97€
 ttcaccatgt tggtcaggct ggtcttgaac tcctgacctc aggtgatcca cccaccttgg
                                                                          977
                                                                          978
 ecteccaaag tactgegatt geaggeatga gecaecatge etggeetgag gaatattttt
 ctaggttccc cccaccccaa gcatttattc tgcaatttta gttttgttcc taaagcaagc
                                                                          978
 aaggtttaag gatttaaaaa taatccgtat tttagaatgc tttctggctt tgttactttt
                                                                          979
 tatccacagt agaagttete agagaatgat etecetettt taatttaaet tittggcaca
                                                                          979
 gtattttgag aattataaat aatattagaa tgttttctgg ctgggtgtgg tggctcatgc
                                                                          980
 ctgtaateet ggetaettgg gaggetgagg caggagaate aettgaaeat gggaggeaga
                                                                          981
 ggttgcagtg agccgaggtc atgccactgc actccagcct gggtgacaga gcaagactct
                                                                          981
 gtotgggaaa aaaaaaaaa aaaaaaagag tgttttottt cotattttoo accacttgat
                                                                          982
 taagttactt ttcctcttaa gtattttttg ctgagtatgc tgacttaaga gtaatgttac
                                                                          982
 aaaatttaat ttttaaagtt ctctgaaagc ccctttatga gagttttagg ctatcaaatt
                                                                          983
 gtgtttaatt ettaacaatt ttttgaaaaa ttatagette aatateegta catteeecae
                                                                          984
 aaaaaagcac taaaaatcat gccttgctgg aggctgcagg accaagtcat gttgcaatca
                                                                          984
                                                                          986
atgccattte tgccaacatg gacteetttt caagtageag gacageeaca ettaagaage
 agccaagcca catggaggcc gctcattttg gtgacctggg taagtaacta tcatttttta ttaacttgta ttagaaggat ttgagtacaa tatgtgaaac ttctgtcata ggatacagaa
                                                                          981
                                                                          986
 ctatataatt ggaaagtgct ttggaaaaaa tgtatttaaa ataacagcta caagtataat
```

```
gagggctgca aacaccagtg accctacaac atttttaaat gtcgtcttcc tgacagcagt gatcagtacc tgcaacgatc ttatttattt ttttcatgtt agtctccaca cacttgaatg
                                                                                    91560
                                                                                    91620
                                                                                    91680
tagaettttt gaaggeaaaa teattgeett ttetgagetg ggageatgte tggeaeatae
caagcactca acagttgatg tattgacttc atccagatac tctgagggcg agttatttcc
                                                                                    91740
tgctactage ctttcacctt tcaatgttta agagcacaaa tacagagatg ggcacgtttt
                                                                                    91800
91860
Caagaaaaga gggttaaaaa tagtottatg toagatootg tgatagaatt cacacttggo
                                                                                    91920
ttaagetget gggeacette etatettgga tgteatatta gettatetae ageagaattt
                                                                                    91980
ttactgtttt atgtagtaag gaagcaatta tatgattatt ttacagacaa attattettt
                                                                                    92040
atotttatt titttagacg gagtototot tigtotocca ggotggagta cagtgtogcg atotoggoto actgcaacot cogootoctg ggitcaagca attototgco toagootocc
                                                                                    92100
                                                                                    92160
aagtagetgg gettacaggt gteegeeace acacecaget cattgttttg tatttttagt
                                                                                    92220
agagatgggg tttcaccatg ttggccaggc tggtcttgag ctactgacct caggtgatcc
                                                                                    92280
accegecttg geateceaaa gtgetggaat tacaggegtg agecacegtg cetggeceag acaaattatt atactetgag tgttagagge ttaggatgtt tteaettgat getatgggag gaataagtaa taagatatga tacacaacca aagacettte tteaetatge ttetagtage
                                                                                    92340
                                                                                    92400
                                                                                    92460
tagtactatg gatgacacat ggtaataata ttggttagca tttgtcctca atttactgtg
                                                                                    92520
ctagttacte ttetaageee ettacaggta tatatttttt tteateaata ateetetaag
                                                                                    92580
gtagttttta ttattgacct aattttataa atcaagaaaa ttaagaccca gagaagtaag
                                                                                    92640
taacttgtcc aagatcacat ggcttataag tggtagagcc agaatttgac cccagatgtt gtgactacat tgtctctcca taagcaggtt caactctttt gactggatgc tgttccaagg
                                                                                    92700
                                                                                    92760
tcacttcctt agagaagect ttgctgacaa ctaccetect gtgccctect ccaaggetgt
                                                                                    92820
ccattgttct agaactitga atactcatct tagaataaag ctggtctaat tittacagtg
                                                                                    92880
ttatagaatg gatctctgac tgcaaaagtt ggtcataatt atcttttat gttctagtga
                                                                                    92940
aaggcaaaga acaagagaag acctcagatg tgaagtccat taaaggtaag ttctgcctt
ggcagtccac tgcattaaaa agtgatgtgc tttgcatttg tgagttcttt aatcctgtta
                                                                                    93000
                                                                                    93060
tactototot titggoatta atcatticig cottattita taattactta tgattitgat
                                                                                    93120
93180
                                                                                    93240
                                                                                    93300
cgcgatettg geteactgea agetetgtet eeegggttea caccattete etgeeteage
ctccccagca gctgggacta caggtgcacg gcgccacgcc tggctaattt tttgtatttt
ttagtagaga cagagtttca ccatgttagc cagtatggtc tcgatctcct gaccttgtga
                                                                                    93360
                                                                                    93420
tecgeeegee teggeeteee aaagtgetgg gattacaage gtgageeace geaceeggee
gtaagtagge tttttttace ttaattttat ttttttgaga tggagtettg etettateee
                                                                                    93480
                                                                                    93540
                                                                                    93600
caggetggag tgcagtggtg ccatetegge teaetgeage atecacetee egggtteaag
cgattetect geetcageet ecegagtage tgggattaca ggtggeegee accatgeeca
                                                                                    93660
gctaattttt gtatttttag tagagacagg gtttcaccgt gttggccagg ccagtctcaa
                                                                                    93720
actoctgace teaagtgate cactegeett ggeeteecaa agteetggga ttacaggegt gagecaccat geetggecat aagtaggett ttactgagee ttgtgtgtat tggetateet
                                                                                    93780
                                                                                    93840
                                                                                    93900
agtgattaca gtgaaccagt gcccttctta ttaatcacac atttaattgt tccctaaaag
                                                                                    93960
tgattagtto actttattta tttagtaaga caaaaaatga agaatactot taactgagca
gtctgttaac tgtaggaaag cactgacact tataaggctt agttttctgt catttatcca
                                                                                    94020
gaagtatggt tgattacagt tittacttit ttattigaat gaacaacctt aatitaaaat atattigtt tattitigt tgggatcgat acattgtcct tgtttataga ttagagcatg
                                                                                    94080
                                                                                    94140
                                                                                    94200
ctttttaaag atgctgtatt actcactgat tttatttgtc cagtgtacag agattgaagt
gggaaaatta taatggaaat tgtttccata gtcattacat attaatttca tcaatttatt
                                                                                    94260
tccataaaat ctgtagattg ctacttattt agatttttcc ttcaaatgtt tttatgttgt
                                                                                    94320
attgcttgca ctgagtattt attctatatg ctcaatttgc tggagaagaa gactaattat aacttaggca agttgtaaaa ttagggaaaa aagtaaggta ccttacagcc tagtttactt
                                                                                    94380
                                                                                    94440
                                                                                    94500
atttettatg taaageeagt tagatteeae attagtteaa aetgeettet ttgageaaaa
cttgattggc agtgataaag gcttaaagcc cttctcaagc agagacctgt aaagactaga tctgactgta gtagaaggaa ggaacttaga tgtttcaggc agtgagaaca ccagtcttcc
                                                                                    94560
                                                                                    94620
actctaaact ttgccactaa cagtatgacc ttgggaagtt gtaactttct tcagattctt catttgttga atggggggat tggcctagct aatttctaaa tctctactgg gctaaaaaat
                                                                                    94680
                                                                                    94740
                                                                                    94800
tctgtgctta tactctgatt atgaagtaca taatctgtgc ttaacattca ctgacttatc
cttaggataa tacagaagca gtacaagaaa cagccctca agatgtttgc agtctggtta
                                                                                    94860
gaaagacaaa cttatacaca gaacagtagc aaatagacca aaataataat agctgccatt
                                                                                    94920
tatagaacac ttettetgtt etgggeatta gacaaaact gactataacg gtgaacaaaa
                                                                                    94980
                                                                                    95040
aagacttagg teetgeeete attgaaetta cagattagta ggggagagga acattaatea
                                                                                    95100
agtaatteca cagatggett agectagatt ggtagtgatg gaagtaaaga gatgtgaacg
```

gacttgaaaa	aaaattcgga	ggcaaaatgg	atagaagttt	attattgatt	aaatatgagg	95160
	agggatattt					95220
22222222	tatatattaa	atttattat	attacataca	agatagttt	gagtgattga	95280
aacayyaaac	tatatgttca	gettegetat	arräddrädä	aggigetete	gagicatica	
	ttatatatgt					95340
ccaaaggtta	gagagtgatt	ttttagagtc	tagcaaacct	aagttgaaat	cctgcctgtt	95400
	tttactagct					95460
	ctaaaatgag					95520
aataatatcc	aaaaaaaaa	aaggtacata	gagaaactcg	tatagtgcct	ggtatatagt	95580
aggtcctcca	ttggtagcta	tcattatcta	gttttaacat	agccttcagt	ttgttgaatt	95640
agtcaaactg	agtgaagcac	tqcaaqqaat	tcagaggaat	ttgagatcaa	caaatqattt	95700
	gggaagactt					95760
						95820
	gaatgagaga					
	gcaaaggggc					95880
atgaaggtat	agaaggctgt	aaagaccaga	aaagagaatt	gagacagagg	aagcaggaag	95940
ccactgtagg	tttttgagca	agatattgat	gctgtaagta	tggtgtttat	gaaaggttag	96000
	atttgcagga					96060
						96120
	ggtgaggtgt					
	caaatcttgg					96180
tttgagtttt	ggagactggt	ggattgaaca	gacagggaaa	ttgagaggag	aatcagatga	96240
tgatgtttta	agttgatatt	tagacagatt	gtgcttgaga	togtaaagtc	aatgtgggtg	96300
	gtagcgagta					96360
					gagctgttgc	
	tgtggagggt					96480
tttgttgttg	ttgttgtttt	ttctgagaca	gattctcgct	ctgttgccca	ggctggaggg	96540
	atcgcgcagc					96600
	cccaagtagc					96660
	tagtagagat					96720
	cagcagagac	999900000	cacgueggee	agaccagacca	tanaaraata	96780
	ctgcctgcct					
cacceggtee	cttgttaagt	ttattttggt	gggaagcaaa	ggaggtttca	gcttttaaaa	96840
agtttgaaaa	ttattgctct	ggtaataatt	aaagatttga	gagtaaatat	gctttctagc	96900
	aaagaagaac					96960
	gggtgttgat					97020
	ggagagtata					97080
						97140
	aagttactag					
	tagtagcatt					97200
tcaccattcc	tatagactct	tacttgtact	tgtctgaaca	cgaaaactgg	cttttgttta	97260
	aaaaattatt					97320
	aaaatagttt					97380
	tgaaattatc					97440
gactactaaa	aatatattt	atgaattttt	agtgtgcttt	atctttttt	gtttttttt	97500
ttgagatgga	gtttcactcc	cgttgctcag	gctggagggc	agtggtgcaa	tctcagctca	97560
ctgcaacctt	cgcctcccag	attcaagcaa	ttctcctgcc	teggteteee	aagtagctgg	97620
gattacagge	acctgccccc	acacccaget	aattttttgt	atttttagta	gagacagggt	97680
treaccates	tggtcaggct	ggtcttgaac	tectgacete	aggtgatcca	cccaccttgg	97740
						97800
	tactgcgatt					
	cccaccccaa					97860
aaggtttaag	gatttaaaaa	taatccgtat	tttagaatgc	tttctggctt	tgttactttt	97920
tatccacagt	agaagttctc	agagaatgat	ctccctcttt	taatttaact	ttttggcaca	97980
gtattttgag	aattataaat	aatattagaa	tattttctaa	ctagatataa	tagctcatac	98040
	ggctacttgg					98100
	ggccacccgg	gaggeegagg	caggagaacc	accegaacac	9994990494	
ggrrgcagrg	agccgaggtc	acgccaccgc	actecageet	gggtgacaga	geaagaetet	98160
	aaaaaaaaa					98220
taagttactt	ttcctcttaa	gtattttttg	ctgagtatgc	tgacttaaga	gtaatgttac	98280
	ttttaaagtt					98340
	cttaacaatt					98400
	taaaaatcat					98460
atoccattte	taccasacat	and to at the	angety-cayy	ascagggga	cttaacaaca	98520
atgccatttc	Lyccaacacq	YACLCCETE	caaytagcag	gacagecaea	cccaagaagc	
agccaagcca						00500
	catggaggcc	gctcattttg	gtgacctggg	taagtaacta	tcatttttta	98580
ttaacttgta	catggaggcc ttagaaggat	gctcattttg ttgagtacaa	gtgacctggg tatgtgaaac	taagtaacta ttctgtcata	tcattttta ggatacagaa	98640
ttaacttgta	catggaggcc	gctcattttg ttgagtacaa	gtgacctggg tatgtgaaac	taagtaacta ttctgtcata	tcattttta ggatacagaa	

```
gggtagctgt gttgtgttcc tgtaaatata gaatataaag catgcccagt agaaaaacaa
gcatttccag aagaaatata tetgatcact aaatataaat atatgaaaaa gatgtetcac
                                                                                        98760
                                                                                        98820
                                                                                        98880
tttattactg agggaagtgc aaattaaaat aatcagttaa tgttctccta acacattagc
                                                                                        98940
atatttttta aagtttgaca atttgaatgt cagtgaagat gcagggaaat acccctccta
tttagtgata atataatctg gtgaagactc tttggaaagc aatttggaaa tcagtataaa
                                                                                        99000
                                                                                        99060
atatgcatgt catttaggcc actetiteta agaectagec etcagatatg etcatteata
99120
                                                                                        99180
tatgtatgta tgtatgttga aggetatica ttatagtatt gtttgtgata geasasaatt
atggacaaca tataaatate tgttataggg aaataaccaa attgtggtat acgcatgete tggagtataa tatagecatt tgtttetatt tatttattt ettgagacag ggttttaete
                                                                                        99240
                                                                                        99300
                                                                                        99360
tgttgcccag gctggagtgc agtggtatga tcatggttca ctgcagcctt cacctcctgg
gcacaagcca ttetetegee teageeteca gagttactag gaetgcagge atgtgteace acacccagat aatttttaa tttttgtag agacagggte teactatgtt geetaagetg
                                                                                        99420
                                                                                        99480
gtotcaaact cotggootca agcaattoto coacacaggo otoccaaagt gotgggatta
                                                                                        99540
ccaacgtgaa ccaccacacc tggttcagtg tagccattta gaaatctaaa aaagacgtgg
                                                                                        99600
gaaaatgtet aaggeatgtt taaatgtgag aaaageaagt cacagtatge atggtaaaat cegttatatt aaaataagtt ettecaaaac aaaaacatat geaggagace tttattttgt
                                                                                        99660
                                                                                        99720
                                                                                        99780
cagtatttet tacccaaatt tetgcactta gaaaattgca tgtcatgttg tcataagttg
aaaaaaagat ccatgaacca atggacttct aataaaatca gtcctgcttt tgacatctct
                                                                                        99840
ctctactttt gtgtatattc aaaccagagt gtcaatgtgt ttgtggggca cacttagcaa
                                                                                        99900
                                                                                        99960
taatacatag cagacaaaat gcatataget cagagagtaa aattgtaagt tttgctagat
                                                                                       100020
cactcataaa ttgctgatga gaatttaaaa tggtgcagat gctctggaaa acaggcagtt
totttottto ttittittit tottttgag acagggtoto actotgitgo gcaggotgga
                                                                                       100080
gtacagtggc gtgattacaa ctcactgcag cctcaccctc ctcaggttca ggtgatcctc cctcagtctc ctgagtagct gggactatag gcatgcacca ccacgcctgg ctaatttttg
                                                                                       100140
                                                                                       100200
                                                                                       100260
tattttttt tttttttt gtagagacgg ggtttcgcca tgtttcccag gctggtctca
aactcctgga atcaagcgat ccacttgcgt aggcctccca aagtgctggg attacgggcg
                                                                                       100320
tgagetactg tgcctggcct aggcagtttg tttgtttgtt tgtttgtttg tttatttatt
                                                                                       100380
tgtagacgga gtctcacagg ctggagtgca gtggcccaat ttttggctca ctgcaacctc
                                                                                       100440
                                                                                       100500
cgcctcccag gttcaagcta ttctcctgcc tcagcctcct gagtagctgg gatgacaggt gcctgccata atgcctggct gatttttgta tatttagtag atatggggtt tcaccatgtt
                                                                                       100560
ggtcaggctg gttttgaact cctgacctca ggtgatcagc ccgcctcggc ctcccaaagt gctgggatta caggcatgag ccgtcatccc tggctggtgg tttcttatga cgtgaaacat
                                                                                       100620
                                                                                       100680
gcaattacca tatgacctag cagttgcact ctgtatttat cccagataaa tgaaaactta
                                                                                       100740
ccttccaata aaaacctgtg cacaaatgtt catagcagct taatattgaa aaactggatg
                                                                                       100800
                                                                                       100860
ttcttcagca ggtgaatgaa ctggttcatt cataccatgg aataccattc agcaataaaa
                                                                                       100920
aggaacaaac tgttgataca tttaaccacc tggatgaata tcaagggaat tatgctgtca
gacaaaaacc agtccctaaa gactacatat agtatgattc cgtttggata atattcttga
                                                                                       100980
aatagagaaa ttaagagaaa tgaaaagatt agtgtttgcc agatgttaga gacagggagg
                                                                                       101040
tgagaggggt aagtgggtgt agttataaaa gtgcaacatg agggatcttt gtgatgttga
                                                                                       101100
                                                                                       101160
agtigiatet tggeagigga igeagaaate teaaigigat aaaattacaa agaactaaaa
acaagaatga gtatagataa aactggggaa atctgaacaa gttagagtgt tgtatcatg tcagtatctt agagtgatat tgtactatag ctttgcaaga tgttaccatg ggagaaacta aagtgtacaa gggatctcta ggtattatta ttttttaga gatggggttt cactatgttc cccaggccgg tcttgaactc ctgggctcta gtgatccgc tgcccagcc tcctaaagta
                                                                                       101220
                                                                                       101280
                                                                                       101340
                                                                                       101400
                                                                                       101460
ctggaattac aggegtgage gaccatgeet ggeeetttea gtattgtate ttagaaette
                                                                                       101520
atgtgaatct agcattatct catagaattt aattaaaaga aattgtaaac ctcacagaag
                                                                                       101580
atcagaattt cctcaagttt gtgatgttga caaagatgaa ctagttgaca ctgacagtaa
gactgaggat gaagacacga cgtgcttcaa aaaaatgatt tgaatatcaa tggattaaga
                                                                                       101640
agaactettt tgacaaattg atgaaaccet cagteagttt tataagaatg cccatettta
                                                                                       101700
                                                                                       101760
tgatcatgct atgaaagcca atttttaaaa aaattttttg tctttcctaa caattagctt
gtggttataa tttaaattta gttaaatata agataaatga ttttttatta agtttagttt catttttcaa ggtacgatct caaagctact ctttaaccta ctatgaatga ataatgctga
                                                                                       101820
                                                                                       101880
gttcataaca totttgtaga tatatocaca attttocoto aggataagtg cotacaagtg gaattactgg actgaaaata atgoagtttg ctaagacttt gotatotgtt cotgaatgot
                                                                                       101940
                                                                                       102000
                                                                                       102060
cctccaaaaa ggttttgcca gtttacatcc tcatgaccag cgaatgagag tgttgcctat
tttcctgtgc ccttgttact gcttaataat ttttgaaaaa aatctaattt gacagacaaa aatgcatttt atgttaattt gcttttctgg gatttttaat gaggttgagt atagttttta atattttat tggccccttt ggaactagta tcataagttt tttttcttaa gaatttatgt agtctgggct gggcgcagtg gctcacgcct gcaatcccag cactttggga ggccgaggtg
                                                                                       102120
                                                                                       102180
                                                                                       102240
                                                                                       102300
```

102360

WO 01/27857

```
ggtggattgc cgaaggtcag gagtttgaga ccatcctgac caacatggtg aaaccgaatc
                                                                                        102420
totactaaaa gtacaaaaac tageteageg tggtggeggg tgeetgtaat ceeagetact
taggaggctg agtcaagaga atcgcttgaa cccgggaggt ggaggttggt tgcattgagc
                                                                                        102480
102540
aaaaaaaaa aaaaaagaat ttacatggtc tgaattgcca ttaaaagaga tatgagaatt
                                                                                        102600
                                                                                        102660
attgagtaac aaataacttt ttaataattt aggcaagttt tggacgattg tactttgttt
agaaaccaaa agcatagtat ttgtagtttt tttatttact ttagttgcta ggaagtaaac
                                                                                        102720
titattcaag gictciggta ccagtigtig ctaaaagtga tigactaatc igicaatcig aaattattig tigcigaact gctaattett tigettetat ettitaggea gatetigtet
                                                                                        102780
                                                                                        102840
                                                                                        102900
ggactaccag actcaagaga ccaaatcaag cctttctaag acccttgaac aagtcttgca
cgacactatt gtcctccctt acttcattca attcatggaa cttcggcgaa tggagcattt ggtgaaattt tggttagagg ctgaaagttt tcattcaaca acttggtcgc gaataagagc
                                                                                        102960
                                                                                        103020
acacagteta aacacagtga agcagagete actggetgag cetgtetete catetaaaaa gcatgaaact acagegtett tittaactga tietettgat aagagattgg aggattetgg
                                                                                        103080
                                                                                        103140
                                                                                        103200
ctcagcacag ttgtttatga ctcattcaga aggaattgac ctgaataata gaactaacag
cactcagaat cacttgctgc tttcccagga atgtgacagt gcccattctc tccgtcttga
                                                                                        103260
                                                                                        103320
aatggccaga gcaggaactc accaagtttc catggaaacc caagaatctt cctctacact
tacagtagec agtagaaata gtecegette tecaetaaaa gaattgteag gaaaactaat gaaaagtgag tatgtgattt tettgtgtgt acatatgtgt eteaetttet tetttaatt
                                                                                        103380
                                                                                        103440
                                                                                        103500
tactaagcag aacttcagat gaggaataaa atgattggaa tattttttt ctcctctaac
tacttgtaaa tttgggagaa tttggagagt gtagtagagt cagatcagtg tatggaaaag gagcaggagt gactggacct tctaagaagt gtgttatcag aattagtaaa tgaagggtca
                                                                                        103560
                                                                                        103620
aatgteetae tttteecete caetgatttt gacateaaae cattateeae atageettat tteeteecte ggtettaatt ttattaatat tttaetgeae tttgeagata aaattttaa
                                                                                        103680
                                                                                        103740
                                                                                        103800
aaaattttta aaaattgcca ataagtgaca tttattaagt tcagtgctta gtgtatattt
                                                                                        103860
ggattttatt tattagtcac aagacctttg tgcaggtagt aggcatgatt atctttttt
ttttgagatg gagtcttgct ctgtcgccca ggctggagtg caatggcgcg gtctcggctc
                                                                                        103920
actgcaacct cogggttcat gocattotec tgcctcagcc toccaaatag ctgggactac aggogootge caccacacco ggotaatttt tttgtatttt tagtagagac ggggtttcac
                                                                                        103980
                                                                                        104040
catgitegec aggatggtet egateteetg actitgtgat eegeetgeet eggeeteeca
                                                                                        104100
aagtgetggg attacaggea tgagecaceg egeeeggaet gattatetta tttacacatg
                                                                                        104160
agaaaaccag ggcttagaaa ggttaggtaa cttcctctag gttgtacagt aaatgtggac
ctagaagcat tttgacaaga gcacctgttt tttttcttc tctattagtt tagaaattat
                                                                                        104220
                                                                                        104280
                                                                                        104340
atactettaa ttateacetg ggattttgat tagacageet teatgttett ttteatetta
aatgttettt gtgtettaaa gggetaagtg atttetteag atettttagt teacteatte
                                                                                        104400
                                                                                        104460
tcagtgaact aaaatgaggt ctaatctgct actgaatcaa gttttcagca tgttatttcc ttcctccctc cctcctcct tccttccctc aaccaggctc ccgaggagct gggattacag
                                                                                        104520
gcgcccgcca ccactcctgg ctaattttta tattttagta gagacggggt ttcaccatgt
                                                                                        104580
tggtcaggct gatcttgaac tectgacete aagtgaceca cetgeetegg ceteccaaag tgctgggatt acaggcatga atcaccacac etgacggcat gttatttca tegcaaagtt
                                                                                        104640
                                                                                        104700
actgtaaget gggagaagtg gcacacaett gtaeteecag ctaeteagga agettaaggt
                                                                                        104760
gagaagattg cttgagccca ggagttttga gaccaacctg ggcaacacag caagacccca
                                                                                        104820
                                                                                        104880
gctcaaacaa agaaaaaaag ttattgaatt ttttatttct atggatcatt ttttgtagtt
                                                                                        104940
tettatteet ticaccette atteceaett ttgateceat ettttattta tttagtttta
                                                                                        105000
ttaaatgtat atttgtctga taattctgct atctacagtt ttttgtggac ctgactcagc
atttettigt tietteggat teagacigtt ggtggetigt gattitagig attitiggee gtgaacatgt tiettggaet tittgtetgig ggaattetet gtgtaetetg tataaattaa
                                                                                        105060
                                                                                        105120
                                                                                         105180
gttacttcag gtgttttgca ttttcttttg ccatgcacct ggggcctggg
                                                                         tcactaccct
                                                                                         105240
totggtacca ottaaaactg aatttttgto ttgggtgete gtactgatee tgtatgagta
caggittata citacigtag aaataiggig titgattaig gggtatigic ccagaiggig
                                                                                         105300
ctggagtatt aatatgctct ctgttaaact taatgtgttg tccctgtaaa actccaaaat tctgaattcc agaatactac tggccccaaa tgtttaagat aagggcactg cctgtatttg
                                                                                         105360
                                                                                         105420
tttctgcctc ccactatttt ccttagttta acacaaactc acctttttaa aaaacatttt
                                                                                         105480
                                                                                         105540
gagagaattc agtattggga agagtttcta acctgtttct ggaaatggaa gtccaaagtc
                                                                                         105600
tgtttctgta attgtttttt ttttgagatg gagtctcact ctgtcaccca ggctggagtg
caatgacgta ctctcagctc actgcaacct ccacctcccg ggttcaagcg attctcttgc ctcagcccc tgagtagctg ggattacagg tgcccaccac catgcctggc tgatttttgaa gagatggggt ttcgccatgt tggccaggct ggtcttgaac tcctgacttt gtgatctgcc cacctcagcc tcccaaagtg ctaggattat gtttctgtaa ttgtaataca tttattgttt ttagaaactg tctttgctt agtggtaatt ttcaataaaa atagaaatag
                                                                                         105660
                                                                                         105720
                                                                                         105780
                                                                                         105840
                                                                                         105900
```

```
cagtggagtt attaaaagag cattagttac atttttccct ttttcattat cttcaaatat
                                                                                                  105960
 tatatatagt aagtttgacc tttttaaaaat gtatacttgt atcagtttta acacatacat
                                                                                                  106020
                                                                                                  106080
 agattectit aactiteace actataaggg taaagaacag ttagtteett cacetttgaa
 gtcaagcccc acctctatcc caacacttgg caaccgctga tettteteeg tetcaatage
                                                                                                  106140
 tttgcctttt ctctttttt ttcttatttt tttttttgag acagcgtctt gctctgtcgc
                                                                                                  106200
 ccgagctgga gtgcagtgag gcaatctcgg ctcactgcaa cctccgcctc ctgggttcaa gcagttctcc tgccttagcc tccctagtag ctgggattat aggcacgcac caccacaccc
                                                                                                  106260
                                                                                                  106320
 ggctgatttt titgtatitt tagtagaaat ggggtttcac catgttggcc aggctggtct
                                                                                                  106380
 caaactettg accteaagtg atceacetge eteggeetee caaagtgetg ggattacagg egtgagecae tgtgeccaat caggaetttt tttttttaaa tttacattea acttgteatt
                                                                                                  106440
                                                                                                  106500
 titticttgt atggattgtg cottcagagt cacacctaag agccctttgc ctaagcaaag gtcatgaaga ttttctcata tgtttccttt taaaagtatt gtggttggcc aggtgccatg
                                                                                                  106560
                                                                                                  106620
 gcttatgcct gtaatctcag cactttgaga agctgaggtg ggcagattac gaggtcagga gatcgagacc atcctggcta atgcggtgaa accccatctc tactaaaaat acaaaaaaa
                                                                                                  106680
                                                                                                  106740
 aaaaaaatta geegggegtg gtggegggea eetgtagtee eagetaettg agaggttgag
                                                                                                  106800
 gcaggagaat agtgtgaacc cgggaggtgg agcttgcagt gagccgagat cgcgccactg cactccagcc tgggcaacac agtgagactc catctcaaaa aaaaaaaaa agtattatgg
                                                                                                  106860
                                                                                                  106920
 ttttacactt tacgtttaga tatatatett ttttgagtta atgtegtata agtatgaggg
                                                                                                  106980
 ttacgtcaga ttttttgttt tttgtttatt tttacatatg gatgtctagt tgttctaata ccatttgttg aaaagacaac ctttactcca ttgaattgcc tttgtacttt tgccatattt
                                                                                                   107040
                                                                                                  107100
 gtctaggcct gtttttggac tcctttttct gtttcatgat gtgtgtgtct attcctttgt
                                                                                                   107160
                                                                                                  107220
 taataccaca togtettaat taetgtatag taagtettaa aattgggtaa tgetggeett
 ataaaacgaa tigggaagtt tttattttta ctcttatttc cattttctag aagagattgt
                                                                                                  107280
 gtagaattgg tgtcatttct tctttagata tttggttgaa ttgggaagtg atgccatctg
                                                                                                   107340
 ggcctagggt titgtttttt gtgtgtgaga cagagtctca cttctgtcac ccaggttgga
                                                                                                   107400
                                                                                                   107460
 gigcagiggt gagatetigg ettactgeaa cetetgeete ceaggiteaa gitateetee
 tgcctcagcc tcccaaatag ctgggattac aagcgtgtgc caccatgccc gactaatttt
                                                                                                   107520
 tgtattttta atgcagacag ggtttcacca tgttagccaa gctggtctcg aacttgtgac ctcaagtgat tagcccacct tggcctcca aagtgttagg attatagatg tgagccaccg tgcctggcag gggcctaggg ttttctttt cagagtatt taaaactatga attcagatta
                                                                                                   107580
                                                                                                   107640
                                                                                                   107700
                                                                                                   107760
 tttaatagat ataggactat ttaagttatc tgtttcttct tgagtgaatt tttactgtag
 tttatggcct ttgagtaatt aattgtattg aattgtcaaa tttatgagcg tgtaattatt tatagcattt cgggtttgta gtggtatccc tcttttattc ctggtgttgg caattgtgtc
                                                                                                   107820
                                                                                                   107880
ttgtttttct ttgtcagatt gtatagggat ttattagtct tttcaaagaa ctagctttg ttttgatttt tctgttgttt tgttttcaat tttattgatt ttctgctctt tattattct
                                                                                                   107940
                                                                                                   108000
 tttctattat ttctgcttgc tttgggttta ttttactctt ttttttttct ccaagttgct
                                                                                                   108060
 taaagtagaa acttagattt ctggtttgag acctttcttt tctaagataa gcatttaata
                                                                                                   108120
 ctgtaaattt ccttctaacc actgctttag ttacaccccc acaaattctg gtattttgaa ctgagcacaa atgaaatgtt ctaatttccc ttgaatctta ttcttttacc aatgaattat ttagaaatat gttatttagt ttgcaagcaa ttggagactt ttttcctgtt atttttctac
                                                                                                   108180
                                                                                                   108240
                                                                                                   108300
                                                                                                   108360
 catttatttc tcatttcatt atattatggt cagagaatat attttgaatg atttcattta
 ttaattttta aaaataacat taaaaaattt tttaaaatgt gaatatacca catacagtat
                                                                                                   108420
 aaagattgta cattotgttt ttggacagtt ttotataaat gtcaagttga tttagttggt
                                                                                                   108480
 taatgatggt gttcagtttt tctttattct tgctgatact ttgtatgcag ttatatcact ttattactca gaagagtgtt gaactttcca actacaattt ttttttccaa ttttacttc
                                                                                                   108540
                                                                                                   108600
                                                                                                   108660
 agetetatet ggttttgett catgtatttt gaggetetgt tgttaggtgt gtacacatte
 aggatgatat cttctgggtg aattgcctgt tttatcatta tgtaattccc tctttatggt aattttcctt gttctaagat cagaaatatc tgttgtccaa tttatataga cactgcagct
                                                                                                   108720
                                                                                                   108780
                                                                                                   108840
  ttcatttgat tagtgettge atggeatate tttttecatt tttttacttt tgatetacet
 ttataattet atttaaagg ggettettgt aggeageata tagttgggta gtgttattta
ttattatt tatttatta tttatttatt tattgagaca gagttttget ettgttgec
aagetggagt geagtggtge aateetgget taceacaace teeaceteet gggttgeagt
gatteteetg ceteageete eeaagtaget gggattaeag geaegegeae eatgeetgge
tgattttttg tattttagt agaaaeggat ttteaceatg ttageeagge tegtettgaa
                                                                                                   108900
                                                                                                   108960
                                                                                                   109020
                                                                                                   109080
                                                                                                   109140
                                                                                                   109200
 ctcctgacct caggtgatcc acctgctttg gcctcccaaa gtgctgggat tacaggcgtg
 agccactgca cccggctgag tcatgttatt tttaatcttt tctcacaata cagggttttt
                                                                                                   109260
 gttggtaaat ttaattattt taatataaat tttagtataa ttatttacat taaatgtaac tgttgcactg gggtatttat aatgtgtaaa tataattatt ggtattaata taattatt
                                                                                                   109320
                                                                                                   109380
 actcataata atattaatat ctttggattt agattaccag tttagtatat gtttttctgt ttctccctct ttgatttccc cttttttgct ttttttttt ttttaattct tatttttt
                                                                                                   109440
                                                                                                   109500
```

```
tagtatttgt tgatcattct tgggtgtttc ttggagaggg ggatttggca gggtcatagg
acaatagttg agggaaggtc agcagataaa catgtgaaca aggtctctgg ttttcctaga
                                                                                            109560
                                                                                            109620
cagaggacco tgcggcctto tgcagtgttt gtgtccctgg gtacttgaga ttagggagtg gtgatgacto ttaacgagca tgctgcctto aagcatctgt ttaacaaagc acatcttgca
                                                                                            109680
                                                                                            109740
                                                                                            109800
ccaccettaa tecatetaae cetgagtggt aatageaeat gttteagaga geagggggtt
gggggtaagg ttatagatta acagcatccc aaggcagaag aatttttctt agtacagaac
                                                                                            109860
aaaatggagt ctcccatgtc tacttctttc tacacagaca cagtaacaat ctgatctctc
                                                                                            109920
                                                                                            109980
tttcttttcc ccacatttcc cccttttcta ttcgacaaaa ctgccatcgt catcatggcc
egtteteaat gagetgttgg gtacacetee cagaeggggt ggeagetggg cagagggget
ecteacttee cagatggge ageegggeag aggegeeece caceteecag aeggggeagt
                                                                                            110040
                                                                                            110100
110160
                                                                                            110220
                                                                                            110280
cggctgccgg gctgaggggc tcctcacttc gcagaccggg cggctgccgg gcggaggggc
tcctcacttc tcagacgggg cggccgggca gagacgctcc tcacctccca gatggggtgg
                                                                                            110340
                                                                                            110400
cggtcgggca gagacactcc tcagttccca gacggggtcg cggccgggca gaggcgctcc
tcccatcca gacggggcgg cggggcagag gtggtcccca catctcagac gatgggctgc
                                                                                            110460
                                                                                            110520
                                                                                            110580
cgggcagaga cactcctcac ttcctagacg ggatggcagc cgggaagagg tgctcctcac
ttcccagacg gggcggccgg tcagaggggc tcctcacatc ccagacgatg ggcggctagg cagagacgct cctcacttc cggacggggt ggcggccggg cagaggctgc aatctcggcactttgggagg ccaaggcagg cggctgggaa gtggaggttg tagggagctg agatcacgcc
                                                                                            110640
                                                                                            110700
                                                                                            110760
                                                                                            110820
actgcactcc agectgggca acattgagca ttgagtgagc gagactccgt ctgcaatcct
ggcacctcgg gaggccgagg caggcagatc actcgcggtc aggagctgga gaccagcccg
                                                                                            110880
gccaacacag cgaaaccccg tetecaccaa aaaatgcaaa aaccagtcag gtgtggcggcggcetgc aatcccaggc actetgcagg etgaggcagg agaatcaggc agggaggttg cagtgagecg agatggeggc agtacagtcc agcetegget ttcacaactt tggtggcatc
                                                                                            110940
                                                                                            111000
                                                                                            111060
111120
                                                                                            111180
gattctcctg ccacagctcc caagtagctg ggactgcagg catgtgccac tacacccagc taatttttt gtattttag tagagacagg gtttcaccat attggccagg ctggtcttga
                                                                                            111240
                                                                                            111300
                                                                                            111360
actettgace tcaagtgate cacetgeete ggeeteecaa agtgetggga ttacaggegt
gagecaceat gecetgeett tttetagaat ttatatattg agttettgat tgtatetttt tatgtagget ttttagtgge ttetetagga attacaatat acataetttt cacagtgtae
                                                                                            111420
                                                                                            111480
tcacatttaa tattttgtaa cttcaagtgg aatgtagaaa acttaaccac cataaaaata
                                                                                            111540
gaactaggga tgaggttaaa aaagagagag aaaagaaatg taataaagat ttaataacac
                                                                                            111600
cgttttttt tittttctc tttttttt gagacagagi ctctctttct gttaccaggc
                                                                                            111660
tggagtgcag tggcgtgatc ttggctcact gcaacctccg cctcctgggt tcaagtgttt
                                                                                            111720
ctectgeete ageetaetga gtagetggga ttacaggtge gegeeaceat geccagetaa
                                                                                            111780
tttttgtatt tttagtagag acggtttgat ttgtgttggc aggatggtct cgatttcttg accttgtgat tcgctctcct cagcctcca aagtgctggg attacaggcg tgagccaccg cgcccggcta agtctttaaa tattttttg acattgcact ttttctcttt tccttctagg
                                                                                            111840
                                                                                            111900
                                                                                            111960
                                                                                            112020
attttagtaa cccaaatgtt agttttgtta ttgtttggca ggttcctgag gctttcctta
cttctttaaa ttttttttc ctgttgttca gcttcgaaaa tttctattca tctgtcttca
                                                                                            112080
aattcactgg ttctttcccg ttatttccat tctgttattg agtctttgta gtgaatttta aattttgtt attatgttt ttagttctaa aattttctt ttttgtgtat gtcttatact
                                                                                            112140
                                                                                            112200
ttgctcctga aactcttatt tgtttcagga gtgatcttat ttcttagagc atggttttag
                                                                                            112260
tagctactta aaatttgttt tatcatccca gcatatgtgt cctcttgatt gtcttttctc
                                                                                            112320
ttgtgagata atgggatttt ctggttcttt atatgacaat taattttgga ttgtatcttg
                                                                                            112380
gacagtttga cttacgttac atgattctga atcttgttta aatcctgtgg aaaatattga agtttttgct ttaacaagca gttgacctag ttaggttcag tccacaaatt ctaagcagca
                                                                                            112440
                                                                                            112500
                                                                                            112560
tictgtegge tetggticea teateagtte agttttgtat ettatetget tatgtgeett
tetgtgteca gtetgggace tggecaatgg teaggtecea aageetttgt acaettttag aageagggee atgeacacee ageteacgag tggeceeggg agtgeacata caactegaeg
                                                                                            112620
                                                                                            112680
tittcatggg ciccitcitt tetgigatgi ceetgacaeg tietgeette taagaacete eetttatee titteetgitg tetggetaga aagteaggge tittagattee etataettea
                                                                                            112740
                                                                                            112800
                                                                                            112860
gcacacttcc tgtagctatg tcaacctctg tggccacgac ttcttcttct tgggactgca
                                                                                            112920
gtttetettg teagaaagta ggattettgg agetgetgte attgetgetg tggetgetet
                                                                                            112980
gatgctgcct gggagtcgaa ggagagaaag gaacaaaaca aaacaaccca ggggatttcc
tecactetet tigatecate agagececet ticetgitee teagaceaga aatagaggge etgeetigga aettetett tigtgeatet gigtgeagtt teagetitig agteeaggee
                                                                                            113040
                                                                                            113100
```

```
aggaggtget ggacaaactt gtcaggagta eggaggtaet gcaagttetg attactttte
                                                                                113160
tragterace tgetteraag teettggatg catttgtera ttgttttgag ttgcattera
                                                                                113220
                                                                                113280
tgggagagac agaagagtgt gettatttea tettgacata ettattagga tttcatatea
                                                                                113340
aatcaacqqa tgatattctc tatattaatt tgctgttttc cctttagcaa gcacattagg
aaataacac tttaacacc gcctttggtg gtttctgtca taattattaa tacttgactt
ttttttttt tttgagacgg agtctcactc tgtcctttga ggcattgtcc ccataaactt
ttggtaaagc atcaataatt ttatcttca tccacacaag cttcaccata aatttgatgt
                                                                                113400
                                                                                113460
                                                                                113520
ttattettee attttageag aatteatgtt geteeaatag gggetgtett caaactgalg tttteteett ettagtgeet cagagtagat eetgtteaga taegttataa eaggttaata
                                                                                113580
                                                                                113640
                                                                                113700
tgagtttatt ttggtgtaaa agtactttga aattcatgca tagttttttc atcatatgca
                                                                                113760
ttttccatag ctttgaacac ccccatgtaa ctctcctctt ccacaaacca aacaatgaaa
aagcaccttt gtgatggaag tttattttgc aataggaact cacagtgatc taagccctgc
                                                                                113820
tattcatgaa tataattcat tactggagtc caagttgctt tttggttttt gaagttctct
                                                                                113880
                                                                                113940
tottcccttg caggtataga acaagatgca gtgaatactt ttaccaaata tatatctcca
gatgotgota aaccaataco aattacagaa gcaatgagaa atgacatcat aggtaagcag
                                                                                114000
                                                                                114060
tgcttgaaac tatggcaaaa aaaaaatgac aaaaaatgca cagaactgac aattttcgtt
                                                                                114120
attgactaag ataatttttt cttaacatgg aatttagcag ttcccttcct aatttgtttt
ctgagtattt tttatatcgg attatagctc actttaaaag tttctcggct gcattcggtg
                                                                                114180
                                                                                114240
cgagggtctt tgcctgggcc agatgggctg cagtgtagcg ggtgctcagg cctgcccgct
gctgagcagc cgggccggcg ggcggctacg ctaaccggca cagaccaccg gatggactgg
                                                                                114300
coggeageec egeaceagtg caegaagtgg gegggacaga aacttetggg gttggaagte cagtgagget aaaageeggt accaaagtet etaggeatea gggetgeage ceaagagtet
                                                                                114360
                                                                                114420
cacgaccagt gggcaactgg atggccagac aggtgtctca gtggtggcct ctccgtctca
                                                                                114480
gggcttcatc ccacttctca gtgggcctga cgtccctggg caccetggat gtctacctgc attagccaga gccatcacat ggcctgtgac ttgccttttt ttgccagttg attgtgccac
                                                                                114540
                                                                                114600
                                                                                114660
acacagtgtc atttctgtgt catttggcac agctggaggt gcaaggagga gggcagcctc
atgtecagte coagtiticae gtaactitat tettetgaat aaagacaatt tgetaacett
                                                                                114720
aaaaaaaaa aaaaaaaaaa agtttttctt atatgttgga cccaaattct taggctttaa
                                                                                114780
cctgaataac aatgacagca agatcaataa atagtacaca tttattaaac actcactgtg
                                                                                114840
toccagacaa tattocaago actititatg gatagactoa tittaactic taaagaacti
                                                                                114900
                                                                                114960
tgtgggataa atacagttat tttatagatg aagaaactga agcacagaga agttaagtgc
115020
                                                                                115080
tatgecetea ecceetgeea aaaaaaaaaa aaaaaaaaaa aaatatatat atatatata
                                                                                115140
atatatatat atatatatat aatatatata tatataaaat atatatatat ataaaatata
                                                                                115200
tgtattagta tatatgcata tatagtatat attatatatt agtatatata ctaatatata
                                                                                115260
atatacatat tagtgtgtgt atatatatat atactagaat aaaaaaatca aagtatctca
                                                                                115320
gagtagtaag gacaaacatt tcagaaaaat gttttcatta tatatacatg tatgtatgtg
                                                                                115380
                                                                                115440
tatgctgatt caacaaatat atttcttata ggttatagca aaatagtttg aaagctttta
                                                                                115500
ctgigtitta tcaggaagac cttaggtgaa cgtatattca cagataaaag aggttattta
ttcattcaat aaatattaca ttctcataag tcctaatatt atgtattttt attcttcaaa
                                                                                115560
aaagttagta tttgtgattt atgaaataag acatgttctt gcacttttag cagatctgtc
                                                                                115620
ccgatgttgg gcttctttaa tccttagtgt gggtgctttg cactcactca ctgctgggga cagcaagace cctgttagte tcagctgtg ttcttaaatt ggcccactgt acettccagt tagctattct ggggtccatg tcatgttgge tccattttcc ttttcttct cccacacaga
                                                                                115680
                                                                                115740
                                                                                115800
tacctataac ggctataaca taggcctggt ggctgttggt ggcttatccc tatctgcttg
                                                                                115860
tatttaaggg gtactgtttc actgagttt gctgacagat gttgtcatga gatttgaggt tttctgtgtt gttgctctat ttttatgtgg gaatttgcta ctatcatcat ccctagacca gcttttccta gtaatacaac agggatgttc tgactgatta gagtttgcct gtttgaagaa ttggttggct agtgatttt ttttgagggg agtctgtacc agttaatagc ctgactggcg
                                                                                115920
                                                                                 115980
                                                                                116040
                                                                                116100
                                                                                116160
tgtggataaa aaggaagcag tttcaagtca aataaaacac ttaaaatgaa accacactgc
aactetettt ettttaetta agettaatea aattaatgat gatgtaatee catgaaggaa
                                                                                116220
aagtettetg aaggateaag tigataacat tttgtgatea aagaatttga gaaaacetet
                                                                                 116280
                                                                                 116340
atcccagtgt ctatcattat atattttagg atgttaatta cctgtgtggc tttaggcaag
tcatttttcc tccttgagcc ccattcttaa tcctgtccaa attatttgtc tcctcttgca
                                                                                 116400
gttggactat tttaatatag ctgtccttca agtgagtttt gttcaaagga gccttcactt
                                                                                 116460
tagefettae tgtgtaecca etttgeatag tettgtttta aatgtaatee ttggattttt
                                                                                 116520
ggtgttgcta actaattact gtttttatgt gaggatttag agtgatccag aatctatact tgcactacct ccttcatctt ccacaaatgt ttgaagtggt agaattttta aaaactttga
                                                                                 116580
                                                                                 116640
                                                                                 116700
aggtacaget gacagaattt getgatggtt tggaagtgag tggtatgaga gggaaaaaaa
```

```
ggaataaagc atgactgcat tttttgtttg tttgtttgtt tgtttttgag acggagtctc actctcgcca ggctggagtg cagtggcgtg atcttggctc acggcaacct ccgcctcctg
                                                                                           116760
                                                                                           116820
                                                                                           116880
ggttcaagcg attcccctgc ctcagcctcc caagtagctg ggactacagg cgctcgccac
                                                                                           116940
cacqcctqqc taattttttt ttttgtattt tagtagaaac ggggtttcac cgtgttggcc
                                                                                           117000
aggatggtet ceateteetg accteatgat etacteacet tggeeteeca aagtgetgag
gttacaggca tatatataag catataaagt gtgttatagc atacaaacag gtatatatat aaacatgcag tccacacagc tgataggaat gaggcagtag tgaaggagaa gttgatgtag
                                                                                           117060
                                                                                           117120
                                                                                           117180
gagaggggac agttgttaca ggaaagaagt ctggaggcag aagggatgaa ttccagtgct
117240
                                                                                           117300
                                                                                           117360
                                                                                           117420
                                                                                            117480
cettetgaac teteacaaaa cagaaceett ceatgactet agttgtgtgg ggttttttee
ctgtcagcta ccaattctgc agatgattgt tcagtgaaca ccaactgggt gtcctctaag
                                                                                           117540
tragttragt totracacty tittacotyga gatagratra gatorcarag attgaggact
                                                                                           117600
ctgtcccaca agactgcctc cacttcagat gccagtctca agtacaagtt gtggcctgtgcttctgactg accttctata aattggagtt cccacagtcc cctccttggg ttcaataaat
                                                                                           117660
                                                                                            117720
                                                                                            117780
ttgctagage ageteteaga acteagggaa atgetttaca tatatttace catttattat
aaaggatatt acaaaggata cagattgaac aggcagatgg aagagatgca tgggcaaggt atgggagagg ggcacagagc ttccatgcac tctccaggtc atgccaccct ccaagaacct
                                                                                            117840
                                                                                            117900
ctacagattt agctattcag aagcccccct ccccattctg tccttttggg ttttttgtgg agacttcatt atataggcat gattgatcat tggctattgg tgatcagctc aaccttcagc
                                                                                            117960
                                                                                            118020
cccttcatcc egggaggttg gtgggtaggg ctgaaagtcc caaacgtgta attetgcett ggtctttctg gtgattagcc ctcatcctaa agctctttag aggccacagc cacaagtcat ctcattagcc ttcaaaagaa tccagagatt ccatgaattt taggcgctgt atgctaagaa
                                                                                            118080
                                                                                            118140
                                                                                            118200
actggctaaa ggccagttgc aatgtctcag gcctgtaatc ccagcacttt gggaggctga ggcaggagga tcgtttcagg ccatgagatc aaaaccagcc tggtcaacat agtgagaccc
                                                                                            118320
                                                                                            118380
ccttacaaaa aatttaaaaa ttggccaggc gtaatagctc ttgtctgtag tctcagctac
tcagaagget gaggatcaet gageeetgga gttgaaggea geagtgagee atgategtge
                                                                                            118440
cactgactcc ggcttgggtg acaaagtgag accttgtctc agaagaaaaa ggaaaaaaaa
                                                                                            118500
                                                                                            118560
aaaactgggc aaagactaaa taacatattt cacagtatca cagatttgta ttgtctagga
aagtgaatgt aaacagacca ggacactagt atgatccctt ggtttcatga aggtcccact
                                                                                            118620
aaagtcatga acacaaagtg agactaggca tcatgttata tggtttttcc agccatgttt
                                                                                            118680
aacagctage taaatageta attgtttege tgeagtttat tttageagtt cettattta geacatttea tgttttaaaa tttetaecaa taacatttta ataaactttt ttacagataa
                                                                                            118740
                                                                                            118800
                                                                                            118860
cttcacaaat ccataatttt ttaagttaca atcccagaaa tagaattgct cattgaaagg
                                                                                            118920
gtatgttcat ttttaaagtt atgctagaaa ctgccaaatt gccttcagaa aaaggtgttt
                                                                                            118980
gtatececae taacactagt gttagtttte ttgtgeeett geteaagtat acatattatt
aaaaacaatg ttgggccagt ttactagata aaaggtgtag tgcctcctta ttctaatcta tttgattact agtgagtatg tatgtcttt cacgttggtc attttatgtt tgttcctttg
                                                                                            119040
                                                                                            119100
tggattgtca tgtcctttgc tcatttttct tttggaacat ttcttagtag tttataagag
                                                                                            119160
                                                                                            119220
ctcttggtat tttaatgata gtaacctttt aactgtcatg catgctgcaa atctttttc
                                                                                            119280
tgtttgtttg cctttgtatt ttgtttttgg agggtttcta tgtataggaa ttaaatttta
tgttgttaaa tcttttgatt tctgcttttg catatgtact tcaaaagact ttctattta agatcaagtg ttacctgtat tttcttttag ttctatttaa aacctcttaa tttatatgcc
                                                                                            119340
                                                                                            119400
                                                                                            119460
tgtgctgtta actcccaagt tgattcacaa gtgtgtatac atagtttgaa tttagtggca
atttaattat ttacaacttc ttttgcagca aggatttgtg gagaagatgg acaggtggat cccaactgtt tcgttttggc acagtccata gtctttagtg caatggagca agagtaagtt
                                                                                            119520
                                                                                            119580
agttcatatt ttcacattgt gcatcctagg gaatttgggt tcattgttag gaatgggctt cactcagcta aaaacaaagt atttttgaga atttaaatat tttggatatt tacaagatca
                                                                                            119640
                                                                                            119700
tataaagcat actctatctt ggttaacagt ttcttttaaa tataaattat gtgaactctt aaaattttca ttttcatttt caatgttaat atttcctaag ttaaaataat ttgttttag
                                                                                            119760
                                                                                            119820
ttctgaaata atttggggag tgattgagtc tgtagtgatt atgactatta gaattggttt
                                                                                            119880
atttatttaa ataatgcatg tetteagatg geteteetaa tetgetaget aggettaag etaaatggat getatataae taaateeaea tagatttgtt gaaatggete eagaggtttt
                                                                                            119940
                                                                                            120000
                                                                                            120060
ttagatttat tactgctatg tgcccttaaa aaaaatctat tcattctttc acttaacatt
tatcàgaaga gtgctctgtg taagacgtgg ttaggcatag tgccagtctt gaaggaagtt acagcctaat aaaagacata gggcatgttg tttggttact gtaatatgaa gtggcatgtg
                                                                                            120120
                                                                                            120180
ttaaatgtca ggggagaact acaaagtcat aaaaaggtgg gagagattac atacaggtaa aggaatcagg aatgacacca tggggagtaa ggtagtgttg acctaggcct ttaagataca
                                                                                            120240
                                                                                            120300
```

```
120360
atagggacag tatggaaaga gtatattttt cccacttaaa ctctttcctt ggtcgttccc
tcaaattttc ccttttqtcc atgtqcaqqc actttagtga gtttctgcga agtcaccatt
                                                                                   120420
tetgtaaata ecagattgaa gtgetgaeca gtggaactgt ttacetgget gacattetet
tetgtgagte agecetett tätttetetg äggtaaagte tgeatttett tteacaetet attegageat tecageetet aactateaat getggggeee tgtetatagg aaataacaea
                                                                                   120540
                                                                                   120600
gaagagccaa gtcatttcca aaaagatgta tcattgtttc aagttgtttc tgatggcaag
                                                                                   120660
                                                                                   120720
agtaatttaa taatatta gagagaacat gaaaattcaa tgtattaaat aactctaatt
ttgagaaacc taattaaact actgcatgta agagagtgca tgtttttaat tatttggagc
                                                                                   120780
tattttaaaa ccacagaatt tgaaacttgc ttccagtgca taaattgcag accagacttc agaagagaaa aaaagtagta aattttttct tatgctcatc atttttactt tagtcacttg
                                                                                   120840
                                                                                   120900
ataqqattqc ccagtgaaga agcatttgca acagacaatg agtatattaa tctttttgag
                                                                                   120960
gcatacagit tagtataatg cictitgita ggcitcaaca agigaaatta tittgitgga aagcaaatga ciattaagta gaaagaggat tcccagictc acaaagcagi aatitagaca
                                                                                   121020
                                                                                   121080
ctogattotg cototttaca agaatacagg tactoagttg atttgtttte teactoott
totttgctat aagtttaaat caacaatttg tttaggttaa tatgtootca tggaatggtg
                                                                                   121140
                                                                                   121200
gaaatgatca gatataaaat atttggtttg gttagtttac tetttatatg tttgctggca aggaaccaca aatccagttt agtataattt ttactctagt tcactaaaag tttgcatcca getgtgtagg tagtgtttgt ttettgttaa ettttttte gtetaaaaga atactttaaa
                                                                                   121260
                                                                                    121320
                                                                                   121380
                                                                                   121440
actiticaat cicaaatgac tgiaactige tgacaggigt taacagaaga agtagaicti
tttqtttttt qcttatqacc tqtattttaa tatttgagct tatagattag agattgtgag
                                                                                   121500
agaaatetgt ttatagiett atttteeett gigtattitt tetteetagt acatggaaaa
                                                                                   121560
agaggatgca gtgaatatet tacaattetg gttggcagca gataaettee agteteaget tgetgecaaa aagggecaat atgatggaca ggaggeacag aatgatgeca tgattttata
                                                                                    121620
                                                                                   121680
                                                                                   121740
tgacaagtga gttatattga tagatggatt cagcagatac ttattgaaca tttgatatgt
tttgtggaaa taaagatgaa taaactcagt ctctgttgtc aaggagctca caggaggcag
                                                                                   121800
cataaaagct gcttttatat ggtgtttgta aagctttggg ggttcttaga acaaaagttt
                                                                                   121860
ctgctgggaa aggggaggtg tatgtggggt aaacaggatg gcaatggtgg tgttcaagga
gtgtttccca gaagagagat tttgtttgga tcccaaagaa agaagggaat tttgctaccc
                                                                                   121920
                                                                                   121980
agagaaggca gaaaacaaca ttctaggcaa aggcattggc ccagaagcca tggaaacgta
                                                                                   122040
ggggaaagtg gcactttcaa gaaacttgag tttagataat caaaggagtg gggaataaat
atgaggatgc tggtactaat tggaatagat tgtaagggac cttgaatgcc tatttatggg
                                                                                   122100
                                                                                   .122160
                                                                                    122220
tatattatac titetgtata aatetgetea ggeaegttgt taattagttt tttattagtt
                                                                                   122280
ttcactgaaa atgagaggat ggaaacatca tacagtaaac aaaattgaaa atatctggtc
aggcagatga tgagettgtg gecagetetg taacgtatgg tattetttte atttaaettt
                                                                                    122340
                                                                                   122400
tottactorg taaaaaaagt aattogtggt ogggcacggt ggctcactoo tgtaatcaca
acactttgag aggcagaggc aggtgaatcg cttgagccca ggaatttgag accagcctgg gcaacatggc aaaacccgcc tttactaaaa atacaaaaat tagctgagcg tgatggcgtg
                                                                                   122460
                                                                                    122520
                                                                                    122580
egectgtigt ectagetaet taggggeetg aggeagaagg ateaectgag eettgggagg
togaggotgo agtgagotgt gatocactgt actocaccet gggcagggca gtagagtgag
                                                                                    122640
accetgeete caaaaaaaa aaaaacaaca aaggtaattt gttattigta teettaagca
                                                                                    122700
aatgctaaag gggtaacttg gggatagaga aaagtccaca gatgttaggg tttgaagaca
                                                                                    122760
ctaatagtat ctaggccagt ggttcctgaa cattagtctg tgggctcttg ctgggctgtc
tgcataggaa tcacctgaga gcttattaaa aataggtttt caggctggtt gcggtggctc
                                                                                    122820
                                                                                   122880
acgectataa teecageact ttgggagget gaggeaggeg gattaettga ggteaggegt
                                                                                    122940
tcaagaccag cctggccaac atggtaaaac cccgtctcta ctaaaaatac aagaattagc
                                                                                    123000
caggcatgat ggcacacac tgtaatccca gctactcagg aggctgagga aggagaattg
                                                                                    123060
                                                                                    123120
ctcgagcccg ggaggtggag gttgcagtga gcggagatca tgccactgca ctccaggctg
123180
ccggtggctc acacctgtaa tcccagcact ttgggaggcc aaggcaggca gatcacttga
                                                                                    123240
ggtcaggagt ttgagaactg cctggccaac atagtgaaac cttgtctcta ctagaaacta caaaaaatta actgggcatt ttgacgggtg cctataatcc cagctactag ggaggctgag
                                                                                    123300
                                                                                    123360
gcaggagaat tgcttgaacc cgggaggcag aggactgcat ctcaaaaaa aaaaaaaaa aaaggtttcc agtccccctg tctcagaaat tctgattctg caggtttgag gtgtgaccag
                                                                                    123420
                                                                                    123480
gaatetttat tittagaaga cataccagat aattetgata aatageeagt ttagggatgt
                                                                                    123540
                                                                                    123600
agtetaattt teetatttig caagtaagga aaataaggee cagagaggta atgattttet
caaagtcaca gaacaagtta gtggcagaat ttggactgga atgcagttct taatgttctg
                                                                                    123660
                                                                                    123720
tccagtgttt attctggtac agtatgtttg tagaaggtat tacgtaagaa acattgttat
                                                                                    123780
atagatgttg agataggaag agtttacatt tagaaatttg gtctaaaatg cctgaacatt
caagtogtgg aggagtattg accaacttac tcaatacaac ataggagatt cacattttgt
                                                                                    123840
tacaaaaatg ctgatttaaa aggagagttt tettttttt ettettttt atttttgag
                                                                                    123900
```

WO 01/27857

```
atggagtett getetgteae eeaggetaga gtgeagtgae aegateteag eteaetgeaa
                                                                         123960
cctccacete ctgggttcaa geggttetee tgcctcagee tectgagtag ctgggattae
                                                                         124020
                                                                         124080
aggtggggg caccacgcc agctaatttt tgtattttta gtagagacag ggtttcacca
tgttggccag gccggtcttg aactcctgac ctcaagtgat ccacccacca ctgcctccca
                                                                         124140
aagtgctggg attataggeg tgagecactg tgeecageet gettgttttt gtateatata
                                                                         124200
tatgcatcat cataatcatg cattatcaac ctttgtattt ctgtcaggac atagaaacca
                                                                         124260
ttagagtgct tggaagagag ccttttttt tttctcgcat ttaatgcttt ttttggtatt
                                                                         124320
cattleataa teagettace aaaacattac etgeattata ecceateaag gtagaaatet
                                                                         124380
ttgtgttatc aatattggtt actocottto cacacogagt catcagtaag tootgttota
                                                                         124440
124500
124560
aagttotoag otogottttt agggaaaatg accatgtott cotttoctat aaattoottt
                                                                          124620
ctatctatca agtoctcaac agagaatagg tacccataaa tatgtgattg ttagtttctt
                                                                          124680
tgcctcagtt gtagtctgat ccttacagct tttaaacaac agtagagttc accgtcaaga
                                                                         124740
actaaggatg gttggcaggc agatagaaag gtagcaagtt gacccaacta tctctgggga
                                                                          124800
agtgggaaca aagaaaggtt acatcagcac tgtcatcaca tagctctata gttctaggcc
                                                                          124860
                                                                         124920
tgcaggctca atcaagtagc cttgtataag attctctgga ggaggtgctg aaagttgctt
atacttgcta tggaatttga ttttacttcg gatatctttt taccataggt acttctccct
                                                                         124980
ccaagccaca catcetettg gatttgatga tgttgtacga ttagaaattg aatccaatat ctgcagggaa ggtgggccac tccccaactg tttcacaact ccattacgtc aggcetggac
                                                                          125040
                                                                          125100
aaccatggag aaggtaacce agaacttcaa acgtatcaaa ctacaagaag ttttattggt
                                                                          125160
agaactcata aaatataagg tgggaaaacc aagcagaata gcacagtgga aattgaagca
                                                                          125220
gtccagcaaa gtgattaaga gcagaggcct tgagtctggc ctggtatgta cagtcacgtg
                                                                          125280
ccacataaca ttttagtcaa cagtggactg cgtgtacgat ggtcctgtac gattataatg gatcaaagct ggtagtgcaa taataacaaa agttagaaaa aataaatttt aataagtaaa
                                                                          125340
                                                                          125400
                                                                          125460
aaagaaaaaa gaaaaactaa aaagataaaa gaataaccaa gaacaaaaca aaaaaaatta
taatggaget gaaaaatete tgttgeetea tatttaetgt actataettt taateattat
                                                                          125520
tttaqaqtqc teettetact tactaagaaa acagttaact gtaaaacagc ttcagacagg
                                                                          125580
teetteagga ggttteeaga aggaggeatt gttateaaag gagatgaegg eteeatgegt gttaetgeec etgaagaeet teeagtgga caagatgtgg aggtgaaaga aagtgttatt
                                                                          125640
                                                                          125700
                                                                          125760
gatgatectg accetgtgta ggettagget aatgtgggtg titgtettag titttaacaa
acaaatttaa aaagaaaaaa aaaattaaaa atagaaaaaa gcttataaaa taaggatata
                                                                          125820
atgaaaatat ttttgtacag ctgtatatgt ttgtgtttta agctgttatg acaacagagt
                                                                          125880
caaaaagcta aaaaaagtaa aacagttaaa aagttacagt aagctaattt attattaaag
                                                                          125940
aaaaaaattt taaataaatt tagtgtagee taagtgtaca gtgtaagtet acagtagtgt
                                                                          126000
acaataatgt gctaggcctt cacattcact taccactcac tcgctgactc acccagagca
                                                                         126060
acttecagte ttgcaagete catteatggt aagtgeeeta tacagatgta ceatttttta
                                                                          126120
                                                                          126180
tottttatac tgtattttta ctgtgccttt tctgtatttg tgtttaaata cacaaattct
taccattgca atagtggcct acgatattca ttatagtaac atgtgataca ggtttgtagc
                                                                          126240
                                                                          126300
ccaaaagcaa taggttgtac catatagcca aggggtgtag taggccatac catctaggtt
                                                                          126360
tgtataagta cactetgtga tgttagcaca atggcaagca geetaaegga aattetgttt
attgattgat tgattgattg attgattgag acagagtttc actccattgt ccaggctgga
                                                                          126420
gtgcagttgc acagtcttgg cacactgcaa cttctgcctc ccaggttcaa ccaattatcc
                                                                          126480
tgcctcatcc tcccaagtag ctgggattac aggcaggcac caccatacct ggctaatttt tgtattttag tagagacagg gtttcaccat tttggccagg ctgttctcga actcctgacc
                                                                          126540
                                                                          126600
                                                                          126660
ttaagtgatc tgcctgcttt ggcctccgaa agtgctggga ttacaggcat gagctaccat
qcctqqqcaq taactqaaat tctctaatgc cattttcctt atctgtaaag tgacgataat
                                                                          126720
atgcacgttt acctcaaagt tactttgatg attaaagtaa ggtaatgtat ataaaataca
                                                                          126780
tattaacata gtacctgaca catggtaagc atcaaaaaat gttaactact tttattacta ttattattac gtatttttaa ataattagag agcagtatca aaaattagct gggcgtagtg
                                                                          126840
                                                                          126900
                                                                          126960
gcatgcacct atagttccag ctactcagga ggctgaagct ggaggattgc atgagcctgg
gaattaaagg ctgcagtgag ccgtgttcat gcccctgcac tccagccttg gtgacagagc aagaccctgt cttgaacaat taaagaaggc attatgccgc aacgttagct tagaaatgat
                                                                          127020
                                                                          127080
ccacatatat caccagtaac tgtcaacagg attggaaccc tagttttggg tattatgatc
                                                                          127140
                                                                          127200
acaaggtatt attaataget tattaataat aaagegttgg etaggeaegg egaeteaeat
                                                                          127260
ctgtaatccc agcactttgg gaggccgagg tgggtggatc acctgaggtc aggagtttga
gaccagootg accaacatgg agaaacooca tototactaa aaatacaaaa ttagoogggo
                                                                          127320
gtggtggtgc atgcctgtaa tcccagctac ttaggaggct gaggcaggaa aatctcttga
                                                                          127380
accegggagg cagaggttge agtgagetga gategeacea ttgcaeteea geetgggeaa caagageaaa acteegtete aaaaatataa ttataataaa taaataaaag taaagtattg
                                                                          127440
                                                                          127500
```

```
127560
atgtttgtga atgatttatt cttctaatga actagaggag atttttccag gaatttcaga
gccagtgagg ttatgttgct tgtatgtgtc atgtgtatcc aggtgaaaaa acttaattaa
                                                                                127620
acgctattat ataataccat acataaaaac tgaattttag gaatactgaa gaatgacata
                                                                                127680
tagaagtcaa atcattaaat agctagtagt aaacagaata gagtgtcagc tgttacccaa
                                                                                127740
tgatgataat attttcacga ttaaaattaa accttttctg attttaaagg aaaagttcag atctgtatca tataaagaat gtaaattttc agggtaataa aattaaaatg cagagagaaa
                                                                                127800
                                                                                127860
aatgcaaaaa tagttottac tagatgtgtg tatgtaagga acttagacta attttaagaa
                                                                                127920
cactgtcaag accetggtag ttaggtagga aaaaagacat gaatgattca ttcaacaaaa actttgagta tttctgtgct agatggtagt gttacagtgg taaacaaaat aaatgtgtt
                                                                                127980
                                                                                128040
ctgctatcct ggagcttagt ctacaaaaaa ggtacatatt ggccgggcac ggtggctcac
                                                                                128100
                                                                                128160
gcctgtaatc ctagcacttt ggaagatcga ggcgggtgga tcacctgagg tcaggagttc
aagaccaget tggccaacat ggcgaaaccc cgtetetact aaaaatacaa aaattaactg
                                                                                128220
ggtgtggtgg cggacacetg taatcccage tactcgggag gctgaggcag gagaatcact
tgaacctggg agacagaggt tccagtgagt cgagatcatg ccactgcatt ccagcccggg
                                                                                128280
                                                                                 128340
                                                                                 128400
ggacaaaagc gaaaatacgt ctcaaaaaaa caaaaacaaa caacaaaggc acgtattaaa
                                                                                 128460
tacgaacata aatatttaca aattatactg aataagttet catgtttatt atttgettgt
ccagttacaa acttttcctt cgtagaatta gaaatataaa taataaacat gagaactcat
                                                                                 128520
tcagtataat taataattat taaatgtaaa taaaaacatc tatgtacaat taggcattta
                                                                                 128580
tttaagaatt atttgaaaaa aaaacaatgt ggaaacagat attttgatat attgctagtg
                                                                                 128640
attgaaattg ataatgttct tttgaagagt aaagtgacca tatatattaa agttaaaatt
                                                                                 128700
                                                                                 128760
taactcagca atcacacgcc tggtgagtta tettaaggaa atcagtttga aagtaaaatc
aatatatgca caaagacttt aacatttatc ataaaccaga aaaatcgagt ttcaaattat
                                                                                128820
atcctatgga ctattttctg ctaaaaagta ttaatatcaa ctttatgtaa tactttcgtg
                                                                                 128880
acaaatattt tgggggagaa aacccaacaa aattacatgc attgtaattt tttttttt
                                                                                 128940
                                                                                 129000
ttttttttta gacagtcttg ctccagogtc caggctggag tgcagtggtg caatctcggc
tcactgcaac ctccatctcc caggttcaag caattctcct gcctcaggcc tcccgagtag ctgggattac aggcgctcac caccatgcct agctaatttt tatagttttt agtagagatg
                                                                                 129060
                                                                                 129120
gggtttcatc atgttggcca ggctggtctt gaactcctgg tctcaagtga tccgtctgcc tcggcctcct agagtgctga gattacaggt gtaagccact gcacccagcc ttatgcatta
                                                                                 129180
                                                                                 129240
                                                                                 129300
taattttaat tigtaaactg tacaaaggga taatacttgt agtacaacaa gaagtaaaaa
catttgttat aggtagttaa catttgtaac cagtagaatt ataggtaaaa tttatttatt
                                                                                 129360
                                                                                 129420
taaaacagtt ttagttggat ttgatttcaa ctitaaaata atgcttttca tctctatcag
gtotttttgc ctggcttttt gtocagcaat otttattata aatatttgaa tgatotcato
                                                                                 129480
catteggtte gaggagatga atttetggge gggaaegtgt egetgaetge teetggetet
                                                                                 129540
gttggccctc ctgatgagtc tcacccaggg agttctgaca gctctgcgtc tcaggtattg
                                                                                 129600
actgattgcg tctgccatta gggagaaaag catacacatc ctttccttca catcccagta acagatccta ttatttgtaa attttaagtt gtggaaaaaa aagataaaag ccaggcacag
                                                                                 129660
                                                                                 129720
tggcctgtgc ctgtaatccc agcactttgg gaggctgcgg tgggcggatc acacgaggtc aggaattcga gaccagcctg gccgacatgg tgaaacccca tctctactaa aaatacaaaa
                                                                                 129780
                                                                                 129840
                                                                                 129900
attageoggg catggtggca ggcacetgta atectageta ettgggagge tgaggcagga
129960
                                                                                 130020
tagoctactt actatottot aatcaaagca tttgtggtaa ottaaaatat actgtattgt
                                                                                 130080
aaagtatcat gctgtttcat ttaggccatt attctatttg aatctgtggc tgtttctctt
                                                                                 130140
                                                                                 130200
aataaatcaa gtaatatgga atatattcat agcctctgaa gagctcttta tgtaagtatt
                                                                                 130260
tatttaggat actttttgta aaataagtga atgaattett aggteteett ttttttett
ttottgagac agggtotoct cgctgcaacc tggaaattot gggctcaaat aatccaccca
                                                                                 130320
ccacagecte etgaataget gggactagag geatgeacea ccaegeetgg etaatttgaa atttttttt ggecaggeat gatggtteae geetgtaate eeageaettt gggagaeega
                                                                                 130380
                                                                                 130440
ggcaggcaga tcacgaggtc gggagatgga gaccagcctg gccaacgtgg tgaaaccccg tctctactaa aaatacaaaa attagctggt tatggtggct catgcctgta atcccagcta
                                                                                 130500
                                                                                 130560
cttgggaggc tgaggcagga gaatggcttc aaccagggag tcggaggttg cagtgagccg
                                                                                 130620
agatcacgee actgeactee tgeatggtga cagagtgaga etceatetea aaaaaaattt
                                                                                 130680
                                                                                 130740
tttttttaaa tgatggagtc ttgctgtgtt gctcaggctg gtcttgaacc cctgacctca
aatgccgcct gcttcagcct aagtttcttt tttttttgta aagagacagg gtcttgctat gttggccagg gtagtctcaa actcctggct tcaagcagtc ctcccacctt ggcctctcaa
                                                                                 130800
                                                                                 130860
agtgctggga ttacaggcgt gaaccactac ctataatgtt gtgtttcact caaggccttt
                                                                                 130920
tgatttegtt ttgcattace gtgccacatt gtgcatttec ttgacetttt ttgggttttttggagtgett tcatatgtta aaccatacet gattetecte aaaatcacac aaagtagaat
                                                                                 130980
                                                                                 131040
                                                                                 131100
atcctaagac aagaaatcta aggaggcata aagaagttaa ctggttttat taaactcaca
```

```
cagtaaatga tagagccaga aatattcccc ttctagtgtt cttcaccatc agcttaatgt
                                                                                        131160
agcataataa ttttctaatt actgttgaca aataaataac cctttgaatt ttcaatactg
                                                                                         131220
ggccttggat aaattttcct aatttgtaag agagtattat cgtattgcca tttacaaagc tctcctgagt atcttttct tctgttaagt ttacctagga gataaactgc tgagtatggt
                                                                                         131280
                                                                                         131340
tgccattttg gttttttgat ataggttaga atgtcttggt ttttttttt tttttttg
                                                                                         131400
gtttttgttg ttgtcattgt ttgagacage atettgetet gtegeceagg etggagtgca
                                                                                         131460
atggcacgat cgtggctcac tgcaacctcc acctcccggg ttcaagcaat tctcctgcct
                                                                                         131520
cagetteetg agtagetggg attacaggea tgtgcaacca cacetggeta atttttgtgt
                                                                                         131580
ttttagtaga gaaggggttt caccatgttg gtcaggctgg tattgaactg ctgacctcat
                                                                                         131640
gatecacetg ceteggeete ceaaagtget gggattgeag geatgageea etgeacetgg etgaatgtet tgtttttgat taggeaceta agaaaggeet aggtactaac cataaaatat atttttatac ettttgttga tactatatat atagaaaacet geacettatea taacettaga
                                                                                         131700
                                                                                         131760
                                                                                         131820
caccttgaag aatgttcaca agcagaacta acccatgtga cccagcatcc agatcaaaaa cagcattatc agccctcta gaagccctct tgggcccctt ccattcactg tccttcttgt
                                                                                         131880
                                                                                         131940
                                                                                         132000
caccagggta getactatee tgacttttga tggcatagat tagcattace tgttettgte
attttataaa taaaaccata cigigiatto tittoiigta cagotttati gigotaatto
                                                                                         132060
acatttacat catacaattc agiggttttt atatggtcac agagttaggt aaccattacc
                                                                                         132120
acatcgattt tagaacattt ttttcactcc agatagaaac cccctttact taaactccaa atcccccact ccaccagecc taggcagcca ctagtctact ttttatctct atagagacaa
                                                                                         132180
                                                                                         132240
                                                                                         132300
tagatttgct tattctggac atttcataaa catggaaccg tatattatgt ggtcttttgt
tgccaactgt ctttcactta gcatcatgtg ttcaaaagag catcatgtta tccatgtttg gcatgtatca gaattttatt cctcattatg gccaaatatc ccattgcaag gatttatgac
                                                                                         132360
                                                                                         132420
attttatttg aattgtaccc teetttetge eatttateaa taatgetaet gtgaceattt
                                                                                         132480
gtgtacaagt ttttgtgtgg atacaggttt tctttttgtt tttaaatttg aggtggagtc
                                                                                         132540
ttgetetgte geccaggetg gagtgeagtg geacaatete ggeteactge aacetetgte tectgggtte aageagttet cetgeeteag ectecegagt atetgggaet ataggeaege aceaceaege ceagetaatt ttttagtaga gatggggttt caccatgttg gecagtetgg
                                                                                         132600
                                                                                         132660
                                                                                         132720
totogaacto ttgacotoaa gtgatocaco catotoggoo toccaaagtg otgggattac aggggtgago cactatgooc ggotgtggtt ttoatttott ttgttgtata tacataggag
                                                                                         132780
                                                                                         132840
tagaattget gagtcaagag gtaactetta aacttattga aaaactgeca gattgtttte
                                                                                         132900
cgaaaagget gcaccatttt gcaateecae cagcagtgta tgagttttac agetteteca
                                                                                         132960
cattreatty gaacttatta tetytttyge tyttttaaa aatgatagte attecaataa gttetaette agtgtygtt ttycaettet etyatyagta atgatyttya geatetttte
                                                                                         133020
                                                                                         133080
atttgcttat tggcctttgt tctagctttg gaaaaatgtt tattcaaatc ctttggccat ttttatttt atttttattt atttatttt ttttgagacc aagtctcact ctgtcagcca
                                                                                         133140
                                                                                         133200
ggctggagta caatggtgtg gtctcagctc actgcaacct ccgcctcctg tgttcaagtg attctcctgc ctcagcctcc cgagtagctg ggattacatt tcaggcacct gccagcatgc
                                                                                         133260
                                                                                         133320
cgggctgatt tttgtatttt tactagtgac agggtttcac catgttagcc aggctggtca
                                                                                         133380
caaactcctg acctcaggtg atctgcctgc ctaggcttcc caaagtgctg ggattacagg
                                                                                         133440
                                                                                         133500
cgtgagccat tgggcccagc ctagattttc ttttttcttt tttttttga gaaggagtct
tgetettgtt geceaggetg gagtgeaatg geacaatett ggeteactge aacetetgee teetgggtte aagegattt cetgeeteag cetececagt agetgggatt acaggtgeet
                                                                                         133560
                                                                                         133620
                                                                                         133680
accaccaca ccagctaact tttgtatttt ttttagagac agggtttcac catgttggcc
aggetggtet caacteetga ceteaggtga tecaectgee ttggeeteec gaagtgetgg
                                                                                         133740
gattaccggc atgagctacc aggcccagcc aattttctca ttatattgcc caggctggtc
                                                                                         133800
tcaaactcct gggttcaagt gatcctcctg ccttggcctc ccaaagtgtg gggagtacag gcgtgagcca ccttgctcag ccctttgcc catttttaaa ttagattgcc tttttatatt
                                                                                         133860
                                                                                         133920
                                                                                         133980
gagtttcagg agtcctttat atattctaga taaatgtccc ttatcaaatt atattatttc
134040
                                                                                         134100
gcagtggcac aatcacaget cactgccace teaactteet gggccgaagt gateetetta etteageete etgaataget agggccatag atacacacta teacacccag ettetett
                                                                                         134160
                                                                                         134220
                                                                                         134280
 ctgtttgtag agacagatet tactgtgttg eccaagttgg teteaaaete taggeteaaa
                                                                                         134340
gtgattetee cacetetgee teccagagtg etgggattae aggtgtgage cacaegeaac
                                                                                         134400
 ctgtetttte actattaata gtgtetteet getteageet eeegagtage tgggattaea
ggcacccacc accatgcctg gctaattttt ttgcattttt agtagagaca gtgtttcacc atgttcaccc ggctggtctt gaactcctga cctcaggtga ttcacctgcc atggcctccc
                                                                                         134460
                                                                                         134520
 aaagtgètgg gattacagge gtgagecact geacceggee aaaatattge ettettaaca
                                                                                         134580
                                                                                         134640
gtattgtctt ctaatttgtg aacatggatg tatcttcatg tatttatgtg ttctttcatt
                                                                                         134700
 tcaqcaqaat tttgtagttt tcagagtaga agcetttcae etecttgggt catttattee
```

```
tatgttttaa gttcttttcg attccattat aaatagaatt gttttcttaa tttcattttc
                                                                                134760
agattgtttg atgagagagc atagaaatac aagtgatttt tacatgttga tcttgcaact
                                                                                134820
                                                                                134880
tcaactttga taaatctgat tgttagctct aatagttttc ttgtggattc tttaggattt
tcaatatata agatcatgtc atttatggat agagatagtt ttttttctgg ctagaactta
                                                                                134940
cagagcaatg atgagtagaa gtggcagaag caaaaatctt tgtcttgttt cctatctgac
                                                                                135000
agggaaaget tteagtttea teatttaata tgatgttagg tgtgggtttt caataaatge
                                                                                135060
                                                                                135120
cttttttcag attcaggaat ttccctatca ttcctgattt tttaaggctt tttttttt
ttaaatcatg aaagggtgtt gaatattgtc atgttctttc tgtatcagta taaatgatcc
                                                                                135180
tatggatttt gggttttatt ctgttgatgt gaaatattaa ttgattttca gatgttaaac caacettgca tacctgagat gaatctcact tggtcatggt gtataatctt ttcaatatgc
                                                                                135240
                                                                                135300
                                                                                135360
tgctggattc catttactgg tattttgttg aagattttgt atctgaacgc ttaagataac
                                                                                135420
atttacacte tatcagaaat gaattgacca taaatgtgag agtgtatttg tgggttettg
attetettee attecaaaga tagacataca teegtetgta tgtetgtett tatgeeagta
                                                                                135480
ccatactete ttgattacta ttgetttgta ataagttitg aaatcagaaa gtataaatga
                                                                                135540
                                                                                135600
gattttggta tetgagtaae agteeteata gaattagttg ggaaatatte eetettatt
etggteete tttettttt gittaactgt gtatetigga gattgtteet teteaacaca tgagageege ttteeetace eteceaece tgetatagag aggtetataa gtgtetgtte
                                                                                135660
                                                                                135720
aattattta tttacttaac ctattactta gtcggggaca ttaagcttgt ttatgtcttt
                                                                                135780
tattttaaac aatgctgcag tgaataatct tgtatataag tcattttcca tcaatataag
                                                                                135840
                                                                                135900
totototgta actgaatttt tagaagtgga atttotaggt caacctatgg ctotgtattt
cacaaaaata ccaattetgg tttttettgt ggaggtgggg agtaggaggt agaatgetgg
aggagaactt getgtaetea getggetagt cattttagaa aggttteett agettettt
                                                                                135960
                                                                                136020
tgtcatatgg cctcaccaag aatcaaaaac attcctattt accctgtaaa catggggctt
                                                                                136080
                                                                                136140
tactacccaa gatacatatt tetggatgta tgacagettt teatattgaa gaaataatge
                                                                                136200
tgtgagtaca gcacatttgt tggaacttag gtcgttaaga atgtcttata aattcataca
                                                                                136260
ttatacattt tattttattt tatttttag tttttgatac agagtcttcc tctgtcgccc
aggecagegt geagtggtac aatettgget caetgegace tecateteet gggeteaagt
                                                                                136320
gatteteatg teteageete cagagtaget atggttacag geatgeacea ceatgeeegg
                                                                                136380
ctaatttttt tatttttagt agaaactggg tttcaccata ttgaccatgc tggcctcgaa
                                                                                136440
ctcttggcct caagtgateg geetgeetea geeteecaaa gtgetgggat cettgtattg
                                                                                136500
ggtaaaagat gaatattgag ggctgcatgg tggctcatac ctgtaatccc agcactttct gagactgagg tgggaggagt cctggagccc aggagggtga ggctgcagtg agttgtgatc gcgccattgc acttcaacct aggaattata ggcttcagtc actgtgcccg gcatgtacat
                                                                                136560
                                                                                136620
                                                                                136680
tttaatattg tgctttcctc ttttagctat agtatgaggt tacatttcag agtcattgtt
                                                                                136740
                                                                                136800
qttaagcatc ttaatagtga tgaggttgag tgaaagttac ttctatttca aacactgaag
aaaattttgt acaaatctgt cacattccaa gcccaggact gattgtttca tatacttcta
                                                                                136860
attttacaat ttctattgta gtccagtgtg aaaaaagcca gtattaaaat actgaaaaat
                                                                                136920
tttgatgaag cgataattgt ggatgcggca agtctggatc cagaatcttt atatcaacgg
                                                                                136980
acatatgeeg ggtaagetta geteatgeet agaattitta caagtgtaaa taaetttgea
tettttaaat titttaatta aattitaeat tittteetaa tetattatta tatgeeeaga
                                                                                137040
                                                                                137100
actttcactt agagtgtgca gtataatgtg gtggttaagt ataaaggctc tggagtgact tcctgggttt taatcttggc tctgccattt attggcagcc gctaacctct tggtatctca
                                                                                137160
                                                                                137220
gtttcttcat ctgtaaaatg agaataataa agtgaaaaga tgccaacatc atttactctg
                                                                                137280
ggctgcataa ctgatacttg gaaaaagtat tcctttgagt ttaagaatta agttggttat
                                                                                137340
                                                                                137400
tcattttagc ttgtaataaa aagatagtga ttcataggat atgccactta ctgaaattta
                                                                                137460
ccacagatec aatcataaaa tcactttete tteeetaaag atagettgat taacatgtaa
aggtgtgtaa aggettgatt acactacect gateegtace ceagtteeca geageaceat gaaaaaggga ttteaacata tttaattact tteagtagaa agtaaeagtg gtaggeeagg
                                                                                137520
                                                                                137580
                                                                                137640
cgcagtggct cacacctgta atcccagcac tttgggaggc cgaggtgggc ggatcacgag
gtcaggagat tgagaccatc ctggctaaca cgatgaaacc ccgtctctac taaaaataca
                                                                                137700
aaaaattago cgggcatggt ggcaggcaco tgtagtccca gctacttggg aggctgagac
                                                                                137760
aggagaatgg cgtgagcccg ggaggcggag cttgcagtga gcttagattg tgccactgca ctccagcctg cgcagtggag cgagactctt gtctcaaaaa aaaagaaagt aacagtggta
                                                                                137820
                                                                                137880
                                                                                137940
ttgggagact gaggagccta gaaagtactt gaaggaagta aaaggtttgt ttgaccacat
                                                                                138000
tgtatttgga aagccagctt tttcagctgt gtcagctttg tgtagtgatt tttagttctt
cttttagaaa ataacggaca aggccgggca cggtggctca cgcctgtaat cccaccactt
                                                                                138060
tgggàggccg agacgggcgg attacctgat ctcaggagtt cgagaccagc ctgggcaaca
tggtgaàacc ccgtctctac taaaatacaa aaagttagcc gggcgtggtg gcgtgtgcct
                                                                                138120
                                                                                138180
                                                                                138240
gtagtcccag ctactccgga ggctgaggca ggagaattgc ttgaacccgg gaggcggagg
ttgcagtgag ccaagatcac accattgcac tgcagcctgc gcgacagagt aagactctgt
                                                                                138300
```

```
ctcaaaaaat aataataaaa taaaaaagaa tggacagtaa acctaaatga gttcattccc
                                                                                          138420
aaagatgatg ttattettaa gggatggtte atttatttaa gacettacat aaagtetate
aattgegtga ttttteactt etgtaattgt gtgtatgtat aatgtaaata tatatgtttt
                                                                                          138480
tgttttgttt tggttttttg agacggagtc tcgctctgtt gctcaggctg gaatgcagtg gtgcaatctc agctctctgc aacctctgtc tcccaggttc aagcgtttct tctgcctcat
                                                                                          138540
                                                                                          138600
cctcccaagt agctgggact acaggcacgt gccaccacgc ccggctaatt ttttgtattt
                                                                                          138660
ttagtagaga tgggggtttca ccgtgttagc caggatggtc tcaatctcct gacctcgtga
                                                                                          138720
tocaccegee ttggetteee aaagtgttge tattacagge atgagecace acacecagea
                                                                                          138780
tgtatttttt aaatgtataa aatgaagcag aaaagagaaa tgataatttt tottoatott
                                                                                          138840
                                                                                          138900
gaaagattat cttcaccagg cgcagtggct cacacttgta atcccagcac tttgggaggc
ctoggoagge ggeteactig agttogaaac cageetggee gacatggtga aacteegtet
                                                                                          138960
ctactaaaaa taaataaata aagatggttt taatatatgt tttagtttta tgattttagc
                                                                                          139020
atotttotga aatttttoto aaggoaagta aatttgtato agttggtata ttggtaccoa
                                                                                          139080
totatgaaat aacttattag gaagatatot otaaaataag atcactttgo otaaaataaa
                                                                                          139140
ctgatatatt gatgttcaca gaatttttct tttaaccgac ttgataaatg cattattctt gacgtcaagt gatccacctt cctcagcctc ccaaagtgct gggattacac acatgagcca
                                                                                          139200
                                                                                          139260
ccgcacctgg cattattett ataaaaggtt aaatttetag ttaagtttaa tgtcctettt
                                                                                          139320
gttcatgtac cattgcttat tttcttccct tcctactcac agtaatcatt cttatggtat gcacttttgt ttgcttattt ttatgtaatt gatattacgc tccattctgt acgttgtact
                                                                                          139380
                                                                                          139440
ttcattcaca gtgagttttg gacattccta tgttcatcta tacagactta cttcatttta actacactgt agtattccgt atgtaatatt tactataact catcactgta gcagagcatc
                                                                                          139500
                                                                                          139560
tcatagtgta tgtattactg ttttgccatt ttggtatcaa tgagtattta agtcatttgc
                                                                                          139620
agtttttccc tcttataccc agtattacag aggatetett tttatatget tetttgtacc
                                                                                          139680
                                                                                          139740
tcactatgtt gcccaggctg gtctcaaact cctaggctca agcaatcctt ccatcttggc ctcccaaagt gctggggtta caggcatgag ccaccatgcc tggcctacat tttaaatttt gatagctctt acaatttact ttgtaaagta tctgcatcat tttatgttct caccagtctt
                                                                                          139800
                                                                                          139860
                                                                                          139920
                                                                                          139980
taataagaat acttcatact tttggctgga cacagtggct cacgcctgta atcccagcac
                                                                                          140040
tttgggaggc cgaggcgggc agatcaagag atcgagacca ccctggccaa tatggtgaaa
ccctgtctct actaaaaata caaaaattag ctgggcgtgg tggcgcaccc gtagtcccag
                                                                                          140100
ctactcgaga ggctgagaca ggagaatcac ttgaacccgg gaggtggagg ttgcagtgaa cttagatcac accactgcac tccagcctag caacagagtg agactctgtc tcaaaaaaaa
                                                                                          140160
                                                                                          140220
aaaagaatac ttcagactta atttttttc cagtcttaag tgtttgctaa tgagattgag tttcttttgg tatgtctctt gattgtcag gtttttctt ttatgaattg actgttcatc
                                                                                          140280
                                                                                          140340
                                                                                          140400
tettetteae attatteetg tegggegatt ttattagtga ettgttaaaa ttetgtatat
tttttcagca tgacacttca ttattcaaaa aaaaaaaaag attctctatg tttctcgata ctatcattg gttggtaata ccttaaaaat aagaccctta ctgtattttt tgctttttt
                                                                                          140460
                                                                                          140520
ttttttttt tttttttt tttgagatag agtcttgctc tgttgcccag gctggagtgc aatggtatga tctcggctct cagctcactg caactgcaac ctctacctcc ctgtttcaag
                                                                                          140580
                                                                                          140640
caatteteet geettageet eccaagtage tgggattaca ggeatecace accacacea
                                                                                          140700
gctaattttt gtattttag tagagacagg gtttcaccat gttggccagg ctggtctcaa actactggce tcaagtgate egectgcete ggcatccaa agtactggga ttacaggcat gagccacagt gcctagccac tttttgcttt ttaactttgt tttatagtac tatagttta gtataaacag atgtatgtat acaccacaact atggcttat aatatgttte agtcattgt
                                                                                          140760
                                                                                          140820
                                                                                          140880
                                                                                          140940
agagcaaggc ctaccttttg ggtgcttctt ttacaaaatt gtcttggcta ttcttgtgcc ttttttctta tttgtgaatt ttagaattgt gaattacctg ttgactcacc atgtttgta aactgaggat tttgaatgga attgcacca attaaagatt atcttgcttt ctgtgcagca
                                                                                          141000
                                                                                          141060
                                                                                          141120
atgttttatt tcaaataatc cctactttaa attacttagg atagctataa attgtgtttc tggctttcta gatttagatg aaacgcttta aattgattgt tttctcctaa atttaaaact
                                                                                          141180
                                                                                          141240
gattgttaga agttaaagte ttetgtteat tettatttag gaagatgaea tttggaagag
                                                                                          141300
tcagtgactt ggggcaattc atccgagaat ctgagcctga acctgatgta aggaaatcaa
                                                                                          141360
aaggtttgtg gtgtttttat acttcatatt aagcctttac tcacattagt gattgactgt
                                                                                          141420
aagtcaaaga ccacttaagg tttaaactgt ttattttgta aagtaaccac tgtatctttc
                                                                                          141480
                                                                                          141540
acettgtgtt tatagteaga agtaagtaea agggetteet gtagteaeat etttatgeaa
totoctotga atcaaaagtt agtgaacttg ctttgccact ccagaaggca catgaatatg
                                                                                          141600
aaaaagcatt gtctattttc ttatttaatg gcaaaatacc cgacctaagt tggacttaat gtttgagacc gtttattta ttaaattata ttttttctct tttcttttt tttttgaga
                                                                                          141660
                                                                                          141720
                                                                                          141780
cagttettge tetgteacee agaceggagt geagtggtet gacegeacet caetgeaace
totgottoot aggttcaago gattttootg cotcatooto otgagtagot gggactacaa
                                                                                          141840
gtgcgcacca ccacacctgg ctaatttttg tatttttagc agagatgagg tttcaccacg
                                                                                          141900
```

```
141960
ttggctaggc tggtctcata ctcctgacct caagcaatcc atccgccttg gcttcccaaa
gtgctgggat tacaagtgtg agccaccatg cctggcctta ttaaattatt tttattaaat
                                                                                              142020
ttcctcaaga ttgatgaaag taatgaaata taaaagtaat gaaatatatg tggaaaatag
                                                                                              142080
actggattaa gaaaatgtgg cacatataca ccatggatac tatgcagcca taaaaaagga
                                                                                              142140
tgagttcatg teetttgtag ggacatggat gaagetggaa accateatte tgagcaaact
                                                                                              142200
                                                                                              142260
gtctcaagga tagaaaacca aacaccgcat gctctcactc ataggtggga attgaacaat
gagaacactt ggacacaggg tggggaacat cacacgctgg ggcctgtcgt ggggtggggg
                                                                                              142320
gctggggag gaatagcatt aggagatata cctaatataa atgacgagtt aatgggtgca gcacaccaac atggtacatg tatacatatg taacaaaagct gcacgttgtg cacatgtacc
                                                                                              142380
                                                                                              142440
                                                                                              142500
ctagaactta aagtataata aatttaaaaa aaataaatat atgtggaaaa tattaatagg
tcaaaattca aattgttcat ttaatcagaa gagtagttta gtcaaatcca agggttagac
                                                                                              142560
aacagaaatc ttttttgtca agtgcattct ttgtgactga tttcattttc ttcctggttt
                                                                                              142620
acacaggaag atttcagaaa caaatgtgga tccgtgacag atggtatcta gaagttttta gtttggttga attgacagta ttttattgag taaaagatac taatttttgt aagaagaaaa
                                                                                              142680
                                                                                              142740
                                                                                              142800
attcaattit gataagtatg tttaagatta agagetattg gecaggeget gtggeteatg
cctgtaatcc tagcactttg ggaagctgga gcaggtgggt cacgaggtca agagattgag accatcctgg ccaacatggt gaaaccctgt ctctactaaa ttagccaggc gtggtggcac
                                                                                              142860
                                                                                              142920
atgeetgtge accegeetee gggtttaage gateetactg ceteaggete etgagtaget gggattacag gegeeatgge taatttttge atttttagta gagacagggt tteactacat
                                                                                              142980
                                                                                              143040
tggccagget ggtetggtet caaacteetg aceteaggtg atetgceege ettageetee
                                                                                              143100
caaagtgctg ggattacagg catgattcac catgtctggc catttatctt attttctttt
tttttttttt ttttgtttga gacggagtct tgctgtgtcg cccagagctg gagtgcaatg
gtgcgatctc agctcactgc aacctctgcc tcctgggttc aagcaattct cctgcctcag
                                                                                              143160
                                                                                              143220
                                                                                              143280
tettecaagt agetgggatt acaggegegt gecaccacat ctagetaatt tttgtatttt
                                                                                              143340
                                                                                              143400
tagtagagac agggtttcac catgttggcc aggctggtct cggaactcct gacctcgtaa
tetgeceace teggeetee adagtgetga gattacaagt gtgagecact gtgcccagec atettattt ettetett ttttgteggg tgggagggg acagagteta getetgeec caggettgge teaetgeace etetgeceec caggetetag caattattet geeteageet eccaagtage tgggattata ggeacetgee acaegeetg getaattett tgttatttt
                                                                                              143460
                                                                                              143520
                                                                                               143580
                                                                                               143640
agtagagatg gggttttget atgttgacca tgctggcctc aagtgatccg cccaccttgg
                                                                                              143700
ecteceaaag taetgggett acaggegtga gettgtattg ggtaaaagaa caatattggg
ggetgcatgg tggtteatae etgtaatetg ageaetttgt gagaetgaga tggaaggagt
                                                                                               143760
                                                                                               143820
gttggagccc aggagggtga ggctgcggct gcagtgaatt gtgatcacgc cattgcactt ccacctaggt aatggagcaa gaccatgtct ctaaaaaaaca aaacacaatt tttttaagga
                                                                                               143880
                                                                                               143940
atactgggaa gaggtcagtg gtggttttag aacagaggaa gtgccagatg acctttgtga
ggcattggcc aggaagaact ctacagtgtc tttaggtagc ttctgtccat aaggataatg
                                                                                               144000
                                                                                               144060
gggteteete eccagtatta atagaaaate tetgagetgt ttttttttgt ttgtttgttt
                                                                                               144120
tgtttttttt teetgagatg gagtetetet etgteggeea ggetggagtg etgtggegatettggete actgeaaget etgeeteea ggtteacace atteteetge eteageetee
                                                                                               144180
                                                                                               144240
caagtagetg ggactacagg tgtccaccac cacgeccage taattttttg ttatttttag
                                                                                               144300
tagagatggg gtttcaccat gtcagccagg atggtctcga tctcctgacc tcgtgatccg
                                                                                               144360
144420
                                                                                               144480
                                                                                               144540
gtcgcccggg ctggagtgta gtggcacgat gtcggctcac tgcaagctct gcctgccagg
tteaagecat tetectgeet cageeteetg agtageaggg accaeaggeg etegeeacea
                                                                                               144600
egeceggeta attitutgia tittitagaag agacggggtt teacegeatt agecaggatg gtetegatet eetgatgieg tgateegee aceteggeet eccaaagtge tgggattaca ggtgtgagee acegtgeetg geetgattit tittitit taatetggte teatacetet gacageteat gaagaagtge teetgettea tatgtatatg tgttageata gtgttaacat
                                                                                               144660
                                                                                               144720
                                                                                               144780
                                                                                               144840
agcataggtg troggtgtt gcagtttotg trigtttat atgaattaag gtgtattatg agcagttgaa gatatatagg aaatttttt ccaaaccact atctetgete gttetattea troagtetgt tratgttatt ccttcattca troatttat agaacagtgg agtgcctact
                                                                                               144900
                                                                                               144960
                                                                                               145020
gtatgcatet attgttetgg gteetgggga agaaaacaaa gtteetgett teatggaact
                                                                                               145080
                                                                                               145140
tacattatat tggcggagac agtaacagac aaacaaatgt agcctgtgta catgtgttac
atgaaaagca gggtagggg ctgggagaga gtagtaggga gtgctatttt cgaggtggtt
                                                                                               145200
gtcaggaaag gcctcactga ggaggtggca ttttgagtag acctgagcgc agcgggggcg taagcccagg cagcatgtgg aggaagagtg ttcttggtga aaggaacaag gatagaggcc
                                                                                               145260
                                                                                               145320
cgaagctaga gagctcagca tgatcaagga acagcaagcc ccgtgtggct ggaatggagt
                                                                                               145380
                                                                                               145440
gagcaaagga atgagcagta gaaggtgagt gagttgggag gtcaccagag accatggcaa
ggacttgaaa gtgtcaggga cacattggaa gttggagcag ggaaatgatg ggatttatgt
                                                                                               145500
```

ttttatgttt	agtgtttta	agggattgct	ctatcagcta	tttggaaaat	145560
gcttcaagaa	gagaagcaga	gaaacaacat		gtcatagtct	145620
tgatggtggt	gtggattagg	ctggtagtgg		cagttcgggt	145680
ggtagaggca	aaaagattat	atttctacca			145740
ttaatttaat	tgagacatgc	ccacataaac	: taataaatag	gaatttctgc	145800
aacacccctg		ttcttcttt	agttgtccag	atgtctcttt	145860
ttttttggtg		ctagagattg	aatttattca	cccaaaaggc	145920
ttactatgtg		tgctgaatgc	: caaggatgta	aataagaggg	145980
gtctgtttta		gttccttttt	aatgaccctg	acttgttaag	146040
atcctacaga			tcctggttgt	.gttatttagc	146100
ttccttgaca			acacctatag	tcttgatgat	146160
cattttagat	_	catgtgttgt		tttgaaagcc	146220
ttggtctgta	_				146280
ttagagatga				taccttttcc	146340
gaactttctt	gccctatgca	tacgttattg			146400
ggctttctgt	tgcatttgga	_		tgttacctgt	146460
tgctggcctc	tatcacctta	ctttgaacca		ggactgagct	146520
tatcttttat	tcttttgctg	_	J J J	cctctgcagt	146580
tggctgtggc	aagccctgcc	atggctttca		gttcctcctt	146640
tattatctct	tcagagaggg				146700
cactcactac		ttttctttc		tttttttt	146760
gagatagggt		tgcccaggct		ctcactgcag	146820
ttgggctcaa		cacctcagcc		ctggaactgc	146880
caccatactt			gtagagacag	ggtttcacca	146940
gtgagccact			tcagcctccc	aaagtattgg	147000
tttacatact			ttcactgtct	aaaattatct	147060
caatactttq		ttatttctca		,	147120
ggcaacaaag	00 00 000	ttggagaatt tctataaaaa			147180
acatgcctgt		acttgggagg	_	aattagctgg	147240
ggttgaggcg		atgattgtgc		agaatcgctt	147300
accatgtgtc	taaaaagtaa		tttctctttc	tagcctagtg	147360
	ggagtttgtc		ggcactatat	atgactagaa	147420 147480
cctagcacat			ttattgactg	ttcctgattc aattatttaa	147540
	gattatctga			atttatgtaa	147600
ttttttaaac			tgacatagac	actaagtttt	147660
ttacttccga	ggatactcac				147720
sttagcactt	tcaagctaat				147780
tctacaaat			acagtggctc		147840
:tgggaggcc	aaggcgagcg		gacaagagtt		147900
gtaaagcct	tgcctctatt	aaaaatacaa	caattagggc		147960
ataatccca		aggccaaggc	aggcagatca		148020
catectgget	aacacagtga		ctactaaaaa	tacaaaaaat	148080
ggtggcacg	cgcttgtagt			ggcaggtgaa	148140
cgggaggcg	gaggttgcaa		ttgcaccgct	gaactccagc	148200
jagggagact	ctgtctcaaa		aaaaacaatt	agccaggcgt	148260
acgagtacct	gtaatcccag	ctactaggga	ggctgaggga	ggagaatcac	148320
jaggtggagt	ttgcagcggg		accactacat	tccagcctgg	148380
jagactctgt	cttaaaaaaa	aaaaaagaa	agaaagaaat	tgaggaatgt	148440
<pre>jtctgtgatt</pre>	tgttaggaat		gttagtagca	actacagggc	148500
lataccacct	tgacaatggt	ttgtttacag	ttcggctccc	cttcctctgc	148560
:tccttattg	agggcagctg	gaaagaattt	tcatcattta	ctagcctata	148620
jagttttgaa	accttgataa			tgagttttct	148680
agtettget		ggctggagtg		atctcagctg	148740
:tgcctccca	ggttcaagca	attctgcctc	agcctctcga	gtagctgaga	148800
jtgtcaccac	gcccagctaa	ttttctgttt	ttatttcatt	ttgtttttt	148860
itcttgctct	gtcacccagg	ctggagtgca	gtggtgcgat	gttggctcac	148920
iteteetggg	ttcaagcaat	tcttctgcct	caqcctcccc	agtagctggg	148980
gtgccacca	tccctagttc	atttttgtat	gtttagtaga	gatggggttt	149040
iccaggctgg	tctcgaactc	ctgatctcag	gtgatctact	cgtctcagtt	149100

```
tcccaaagtg ctgggattat tggcacacgc ctatttttgt atttttagta gagacggggt
ttcaccatgt tggttagact ggtctcaaac ttctgacctc aagtgatttg cccgcccag
                                                                                                    149160
                                                                                                    149220
cctcccaaag tgctgggatt acaggcgtga gccaccgtgc ccagccaaga ttgagttttg aaaagagcct tctgagatta tgagaaggc aagcaagata acttaagaag ttacattaaa
                                                                                                    149280
                                                                                                    149340
                                                                                                    149400
atcatctaag agacagtgta acaagaagga attgtaaaat gatgttatga gcacgtgccc
aatgtagtgg caatcccttg tgcttcgata cattggtggg agacaaaact gtacttaaat tgataaatcc cttacatgtc attttaagga gcttagactg actcccatca tgtagacatc
                                                                                                    149460
                                                                                                    149520
149580
                                                                                                    149640
ttgttgccga ggctggagtg caatggcgtg atctcggctc accacaacct ccacctccca
ggttcaagca attetectge etcageetee egagtagetg ggattacage catgeaceae caegeetgge taatttgta tttttagtag agaeggggtt tetecatgtt gtggetggte
                                                                                                    149700
                                                                                                    149760
tegaacteet gaceteaggt gateeteeg ceteageeac ceaaagttet gaaattacag
                                                                                                    149820
                                                                                                    149880
gegtgageca eegegeecag eccagagatt tetaaacaga gttetaacca gatgetttte
cctgtcagta gaatgagaat gaattggagg tgggagagac tggcatgagg gacaccagtc
                                                                                                    149940
agccagtgga attagctggt aatgttgata ggagaagaaa aagattcaaa gttaggtagt ggtagcaaga attagaggga aggtcggatt tatgatatgt ccaaggttga attctaaggt gaaatttggt ggcagatttc atgtgtaaat tgggaaggta gattgagttt ttttaacatg
                                                                                                    150000
                                                                                                   150060
                                                                                                    150120
ggttttctaa catgtcaata gagtgactct gcaggggggc ctgacgagag aacagtgcat
                                                                                                    150180
ggggtgattc aacagccagt tgagccttca tgcagagcat ttaacactgt gactctgtag
                                                                                                    150240
actotggttg goagtaaaat ttoattaaac caatatttaa accottaggt aataataaaa
                                                                                                    150300
attgagggaa aaggatccag gttttgtatt ttttatgaat tcagttattg aattaaacag gacettgeet caagaaataa tetaecaaca attaaettgt tttaaagcaa agttaggaag
                                                                                                    150360
                                                                                                    150420
                                                                                                    150480
tgagcatgtt caaattatta aataaaaaag taagctgtgt atttcattca tagaaataga
ggctggccta cttcggatga ttctcagcat gtgattacag atgtgggctt atacatccta gggagttaag gcgtactctg gcttggatag agtagagctc tttgaaactc ttctctcacc
                                                                                                    150540
                                                                                                    150600
cagctagttt atatagacta gagaactaga atgtagcagc atactctgtc ttagaagccc
                                                                                                    150660
tittatatag gagctggtct ggaaggtttg aaaacataac aaatgtgttg gtgtctccca atgtattgct agattcttac ccaagagcat tatcctggtt agggtttggt ttggttttgt
                                                                                                    150720
                                                                                                    150780
ttigtttttt aatgtttgcc acaaactaac actagatgtt agttctttca tcaagtgagg
                                                                                                    150840
agagtagaag aaaagtccag aactctgaaa caccittica aaagtttttc aagccatgat
                                                                                                    150900
gtttgcaagt taaatgctct gttatgtaag caatataatc agtttttatt aatgtaacat
                                                                                                     150960
teettagtgt titggggtat cacacaaaa agaatateea tatetggaag caacagettt taaataagag cattgtggtg gtggtggtga tagtggtttt tittittit titgagttgg
                                                                                                    151020
                                                                                                     151080
agtotogoto tgttgcccag gttggagtgc agtggcacga tctcagctcg cttcaacctc tgctcccagg ttcaagcaat tcttctgcct cagcctcctg agtagctggg attataggca
                                                                                                    151140
                                                                                                    151200
cctgctacca tgcctggctg atttttatta ttttagtaga gacaggtttc accatgttggccaggctggt cttgaactct taacctcagg tgaatcaccc acctcggcct cccaaagtgc
                                                                                                     151260
                                                                                                     151320
tggaattaca ggcatgaacc accatggcca gccaaataag agcattttta atgtaaaatt atgcatgaaa tgtacattca attttgtctt tgtttactag gatccatgtt ctcacaagct atgaagaaat gggtgcaagg aaatactgat gaggtaaatc ctacctttag gataaaaaga
                                                                                                    151380
                                                                                                     151440
                                                                                                     151500
tttctgttta taagtgccac cctcatgtaa gtgaggttta aaattttcct tttctttagg
                                                                                                     151560
toccatgttt aagcagcatg gcacatttat gttctcttac ccagaatgta ccaagaaagg gtggtccctt cttaacatct aacaattgcc tggtagtagc agtgaaggta tcttcagtca
                                                                                                     151620
                                                                                                     151680
gaggetagga ceaetgaagg atatacatge atteaagttt ceateageca geaggeatea gtaateagtg tgtagateaa aageteaaat gttteettee ceaetggeag ttttaettea
                                                                                                     151740
                                                                                                     151800
agtagtggag gcttgcttt ttaatagtta attaagtaca ttgagagatg ggaggtgaaa aaaggaaaat gttttatttt gaccatctaa tatgaaagta gttcggtgtt aggtatccag
                                                                                                     151860
                                                                                                     151920
tagttgacac tggaagacag ggaatgacat gttaatattc atagccagag ggtggcccag
gtttttcgt acatgggaat gaaattctta tccaaataag tagaaattat gtgcgtaagc
cattgttaa gagcactgag tatgtgcatc tcgatccatc taatgaataa ccattatcac
                                                                                                     151980
                                                                                                     152040
                                                                                                    152100
                                                                                                     152160
cagtttaaat tattttcttt aggcccagga agagctagct tggaagattg ctaaaatgat
agicagigae attaigeage aggeteagta igaicaaceg tiagagaaai ciacaaaggi
                                                                                                     152220
aaggatgact togttitgig taaactaaaa agtattattt tocaggigta aaaataaaaa
                                                                                                     152280
agaacataag gggtttettt geetttgaag gattaactge tgtggggatt acettettat cataagcaac tagaaaattg acaaactaaa tgaaacaact gtttgcatat attggacaat
                                                                                                     152340
                                                                                                     152400
gggcaataca gggaaaccat ggaaaccaaa cagagcccag tagtettget gaaegaaaga
gttaaatate aaagtteagg ceaggtgeag tggeteaege etgtaateee ageaetttgg
                                                                                                     152460
                                                                                                     152520
gaggecaagg egggtgaate aettgaggte aggagtteaa gaccageetg gecaacatgg
                                                                                                     152580
tgaaaccctg tcttagccgg gtgtggtggc aggcacctgt aatcccaact atttgggagg
ctgaggcagg agaatcgctt gaaccaggga ggcggaggtt gcagtgagcc gagatcacac
                                                                                                     152640
```

```
cactgcactc cagcctgggc gacgagcgaa accccatttc aaaaaaaaa tcaaagttca
                                                                                   152760
gagageteaa titgagtaga agtigtagga taaggtagea gaaaagagga agetgeeeag
                                                                                   152820
aaaqaaagcc gtagagatat ttagagagat toccatggat cettggeeta ggagtgatet
                                                                                   152880
gtatatgtgt ggggtgaaaa cgcatgtgtc caggtagaga acccccaga aattagtagg
                                                                                   152940
ctgaatgatt gctggaacat agggctaaga aaagttcatg gccagaagga tctggccaga gtagagagac ttagtaatac acaaggcatt gggtagtgtc ttcacagagg ttatgcctta
                                                                                   153000
                                                                                   153060
ctactgaaga taaattagto otagagtaca agcacctgaa ccaagtttoa aagcaaattt
                                                                                   153120
                                                                                   153180
ttaaagggtc aaattaccta acaactgcat gccaaaacaa aggcctaacc ctctttacag
taacacaaca aaattcagca cttcacagtg taaagttaga atgtctgacg tccaggctgg
                                                                                   153240
gcgcagtggc tcatgcctgt aatcccagca ctttgggagg ccgaggcagg tagatgacct
gaggtcagga gttcaagacc agcctggcta acatggtgca accccgtctc tattaaaaat
                                                                                    153300
                                                                                   153360
acaaaaactt agccaggcat ggtggeegge acetgtgate eeggetaett gggaggetga
                                                                                   153420
153480
                                                                                   153540
cgtcaatcac aaattaccaa gcatgacatg aagttgacct ataaccagga gaaaactcaa tctatagaaa cagacccaga tgtgagaaag atgatgaatt tagcagacaa agaccatcaa
                                                                                   153600
                                                                                   153660
gtggctattt taaatattaa aaatatgttc aagtggccag gtgcagtggc tcatgcctgt
                                                                                   153720
aatcccagca ctttgggagg ccaaggtggg taggagttca agaccagctt ggccaatatg
                                                                                   153780
gtgaaacccc ttctctacta aaaatacaaa aaaattagct gggcatggtg gcaggtgcct
                                                                                    153840
                                                                                   153900
atagteecag etatatggga ggetgaggea caagaateae ttgaaceegg gaggtggagg
ttgaggttgc agtaagccga gattgtgcca cttgtactcc agcctggaca acagagtgag
                                                                                   153960
actotytoto aaaaaaaaaa aaaaaaaagt taaagaaaac aagagtataa tyagaaaaat
                                                                                    154020
gcaaaatagt tttaaaagaa ccaaatggaa tttcttaaaa taaaaaatac cagaaatggg
                                                                                    154080
                                                                                   154140
ggccgggcgt ggtagctcac gtctataatc ccagcacttt gtgggggctg aggcaggcag
atcacctgag atcggtagtt caaggccagc ctgaccaaca tggagaaacc tcatctctac
                                                                                    154200
taaaaataca aaattagctg ggcgtggtgg cgcattgcct gtaatcccag ctacttggga ggctgaggca ggagaattgc ttgaacccgg gaggcagagg ttgcggtgag ctgagattgc
                                                                                    154260
                                                                                    154320
accagtgcac tccagcttgg gccacaagag tgaaactccg tctcaaaaaa aaaacaaaaa aaaacagtag actcgaagaa ctagctgagt ttttctttac tttaggcagt aagtgtgacc
                                                                                    154380
                                                                                   154440
                                                                                   154500
ttttgcaggt gactacttta gttcctcatg tcctcattag tagatcagag aaattcgaca
ccaaaacccc aaaagaaaaa ccccttctaa tcctcattcc atgattttat gaatgcatga
                                                                                    154560
agtectagge etgegaagga atacteatte tetttateet gtgttgatae etetetgett
                                                                                    154620
caacetecaa etegacattt geetatagga tgtaettgga catteageat aaactacete
                                                                                    154680
                                                                                    154740
acaccattac tgaattgett catgtgeaca tgtcccatgc cacaataccg gggaccttgt
                                                                                    154800
cttccgtgat atttgtccgc agtgctgtga ctacaggagg gagtcagtga atgtctgcat
gtgtgtcttt accatecete ttgaatatge tetagggtta attectagaa gtagaattae
                                                                                    154860
totattgaaa attggcaata ttittcattc taatatctat tgccaacatg ggaaagcaag
                                                                                    154920
totggatgcc agtccttgtt atatgcccct tgggtaagtt acgtaacctc tttaagcttc
                                                                                    154980
                                                                                    155040
tgttcactca tattttaaca aggaaaatta caatatttta cctcacaaaa ttgtagtcag
cttctggctg tcttaaactc tggtatatag taaacactaa gtgttggtgt ccatccttaa
                                                                                    155100
tttgtaataa taggtcactt gttagagaaa tgcaccttac cattttcttt tctttcttt tttcagttat gactcaaaac ttgagataaa ggaaatctgc ttgtgaaaaa taagagaact
                                                                                    155160
                                                                                    155220
                                                                                    155280
tttttccctt ggttggattc, ttcaacacag ccaatgaaaa cagcactata tttctgatct
                                                                                    155340
gtcactgttg tttccaggag agaatgggag acaatcctag acttccacca taatgcagtt
                                                                                    155400
acctgtaggc ataattgatg cacatgatgt tcacacagtg agagtcttaa agatacaaaa
tggtattgtt tacattacta gaaaattatt agttttccaa tggcaataac ccatttatga gagtgtttta gcctactgga atagacaggg accacatcct ctgggaagca gataagcata
                                                                                    155460
                                                                                    155520
                                                                                    155580
gaactgatac ttgatgcaca ctcgtagtgg taactcatcc ctaatcagca ttgtaaagca
                                                                                    155640
ggtgccagag gtggtttgct ttgtccttcc aaagcaggtg agtcagcccc accgagagcc
aggcagettt gagtggcage gtggtgetag cagetteage ggaacagggt gagagttaat
                                                                                    155700
tatgcagtct tottgacage ggcattaatt tggaaggaaa ctgacaagtc atgggtcaag tttcagtgac ttcctccttc ctctgatggc agtatatagt tttcacattt taattcctcc
                                                                                    155760
                                                                                    155820
tectgagatg cactataett aaaaceatte teteceetge taacagaagg gtgtgaatet ggtttaettt gageattagg atttgeeeet ttggaattet geacteeagt taettaaett
                                                                                    155880
                                                                                    155940
tcccttcaga atacatgtgg aaagaaagaa agaaatagcg atgactcac ttttgcccct gtggcacctt gaacaaagca gttcttccca aattatactt tttttttt taaataaggt gagcaggatg actggggaga gagaaacatt tgactttgac tgcctcccc attcttgct gtggcctgga aagtgtgcag ttggtcgtct ttcttctcct ttctttagga tagtaagaga
                                                                                    156000
                                                                                    156060
                                                                                    156120
                                                                                    156180
ctcactcact geacttetge teagttgget tetgeategg gateacacag ccateageag gaetgeecag ttggtgagea cactceattg accaegegge gecagegett ceteaatgea
                                                                                    156240
                                                                                    156300
```

```
catgattgag aggaaagaaa gttctcttag atgttactgc ttttgctcag actttgcaaa
                                                                                   156360
aaaaaaaata tatatatata tgtataaata tataattatt aatcactttt gtccttgaga
                                                                                   156420
aagtottgaa tgaacagaga atttattoca ttgcaatatt tgattgtata gaggcacact
                                                                                   156480
gtttcatcga cagaagaagc aaaaaggctt tgtgtaagtt tttggtacta tgtaccacct
                                                                                   156540
                                                                                   156600
ctgttattct tttaaagetg aagtatteat gtaettaaae catattatat ttaattgtgt
ttgattttaa aatatatata tatgaattet atttaaaatt gtgtcaactt tetgetttea
                                                                                   156660
gggcatttat ggctcttctg ttgaaatata ttgatctttc caaatatttt catttgcttt
                                                                                   156720
ctaaaaaccc agaacatgag ccactactgg actttgcctt gtgtttgaag tgtatggcat aaacccaagg tttttattag tcatctatgc tgtgattaat tcattttgtt cttttaacaa
                                                                                   156780
                                                                                   156840
                                                                                   156900
aatatttcca tccacttcac attgcttcaa tctttaacag aaaagcaata taaaggttat
agaataaaat gtggttttgg gcaactcttg ctgcctctgc atgttttgga ataacaattt ctacaagact ctaggctgtt taaactagtg ctttcagtta agataaattc taatcatttc
                                                                                   156960
                                                                                   157020
titgtatata cattitgige tietgageta gagatgecaa giagttgtaa aetgettata aagagaatag cagcaaatti gagaetegge taettitte tgeeceacet getitgagae
                                                                                   157080
                                                                                   157140
acagaagegg agtgtggeee gaaattatta gecagattta atatttgate taaagtaggt
                                                                                   157200
cettgtacte attttaaagt tggaatttga tteetecaae attgageace caccatgtte
                                                                                   157260
caggetetgt geattgtgee cacaaaataa gatteeetgg tggagttttt atgggtteaa
                                                                                   157320
ataatcagtt gaacaccctt catctttatc atgttgttga cattgacaca aattgtttaa
                                                                                   157380
aaagaaaaga tattagagag aaagtggtac ctttgtaact tgatgtgtct tcatcattcg
                                                                                   157440
gtaagatttg atgaaagtaa aaagcaaatg tcagccaaat ccagtgaaca gcaataaaac
                                                                                   157500
agggagtaac ttittataac tttitctact tggatttcaa caticagtag agcttttcga
                                                                                   157560
aatgtaagta gtttacagta ctggaggttt gactagttca gtaggaattt ggaggggaag
                                                                                   157620
gtcattctga attgtaacaa agtacaaact tctttgctgt tttatttaag tactgagage
                                                                                   157680
                                                                                   157740
taagcacctg atgaagtgac tgacetetet ceagtgacag tgtttgggta cetgeetgae
ttcaggagtg gggtttatgt ttctacacag tgaccttttc tctcgccctc tcctccctct
                                                                                   157800
tgcccacaca ccagttgatt ggacctgggt tgaactcctg atccagacag gcccaagaca gttcttaatg ttaagaattt tggggccggg cacggtggct catgcctgta attgcaacac
                                                                                   157860
                                                                                   157920
tttgggagge egagacagge ggatcaettg aggtcagggg ttegaggeca geetggecaa catggtgaaa ceetgtettt actaaaaata caaaaattag etgggeatgg tggegeaege
                                                                                   157980
                                                                                   158040
ctgtaatccc agctacgtgg gtggctgaga cagggggaatc gcttgaacct ggaggcggag gttgtgcaat gagccgagac cgtgtcactg cattccagcc tgggtgacag agggagactc
                                                                                   158100
                                                                                   158160
tgtotocaaa aataaaaata agaaaaagaa ttttgggota ggtgcagtgg ctcacgcctg
taattacago attttggaag goccaagatg ggcagatcac ttgaggacag gagttcgaga
                                                                                   158220
                                                                                   158280
ccagcetgga caacatggtg aaactecate tetactaaaa agacaaaagt tagecagatg
                                                                                   158340
tggtgatggg cacctataat cctagctcct cgggaggctg gggcaggaga atcacttgaa cccaggaagc agagattgca gtgagccaag atcacatctc tgcactccag cctgggcaac
                                                                                   158400
                                                                                   158460
                                                                                   158520
agagcaagac tetgteteaa aaaaaaaaga atttggeeag gegeagtggt teaegeetgt
aatcccagca ctttgggagg ccaaggcagg cagatcacga ggtcaggaga tcgagattgt
                                                                                   158580
cctggctaac atggtgaaac cctgtctcta ctaaaaatac aaaacattag ccgggtgtgg
                                                                                   158640
tggtgggcac ctgtagtccc agctactagg gaggctgagg cagaggaagg atgtgaaccc aggaggcgga gcttgcagta agccaagatc gtgccactgc actacagtct gggcgacaga gtgagactcc gtctcaaaaa aaaaaagaat tttggccggg tgcggtggca catgcctgta
                                                                                   158700
                                                                                   158760
                                                                                   158820
                                                                                   158880
gtoccageae tttgggagae caaagtggge ggattacetg aggteaggag ttcaagacea
gtccggccaa tatggcgaaa ccctgtctct tactaaaaaa aatacaaaaa ttagccaggt
                                                                                   158940
gtggtggcgg gcacctgggg aggctgaggc agggagaaat gcttgaaccg gggaggcaga ggttgcagta agccaagatc gtgccactgc actccagagc aagactcttt ctcaaaaaaa
                                                                                   159000
                                                                                   159060
aaaaaaaag aattttgcat ggggaaggag agatactgtt caccatctgg aatggtgctt
                                                                                   159120
                                                                                   159180
ggatgtggca cttacaaaat caggagccag cactgcatgg acaaacagaa gcatgtgggc
ctgagatage aggtacettg ataaccetga agacateett ggtttetgea tetatteetg
                                                                                   159240
catcettgea tiggactaca ttaatetgic agttateett ataatgattt tigattitt
                                                                                   159300
ttttttgaga tggagtttcg ctcttgttgc ccaggetgga gtgcaatggc acgatetcgg ctcaccacaa cetecacete ccaggttcaa gtgattctgc tgcctcagec tectgagtaa
                                                                                   159360
                                                                                   159420
ctgggattac aggcatgegc caccacacct ggctaatttt gtatttttag tagagacggg
                                                                                   159480
gtttctccat gttggtcagg ctggtctcga actcccaacc tcaggtgatc accctgtctc
                                                                                   159540
ggcctcccaa agtgctggga ttacaggcgt aagccatggt acccggtctg ttttttgatt
                                                                                   159600
ttttgaaacc agtctgaagt gagtttttt aattacgtga aaggagtttg gctaaaatac
                                                                                   159660
tgccatactg ccctaatgcc taatgattat gtattetcag catgtetgca aagtactgct
                                                                                   159720
                                                                                   159780
gatttetgga gaataatttt tetttagtaa aetteaetta agtegteatg tgtattetet
caaaatggta teetaaceta atggagetaa aagacaceee ttgtttttat aacaagcagt
                                                                                   159840
tactgaggcc caggaagggg agaagtccct ggcttgtgag atgatcacca ttagaactca
```

```
ggcctgggcc agtgcctttt catgcttctc agatccttcc aaagaataat gaagattata
accectttta gcaattetaa taaacccaga aatagaaagc tttttegetta gagtactegt
                                                                                  160020
agaagtttgg cgggagagat aatttttaca aaatttgtaa atacctgcca attctatata
                                                                                  160080
ctaggcaagg tototggoot tgtaaaaccc ctcaaggtta caactttggt ggcccacact
                                                                                  160140
                                                                                  160200
aatagttacc cactgaggcc ctctccgggt gaacattgag cactagagga agcccctctg
cttgggcagg actgggcgtg gtgcagagta ggagcggtga tactgtggat tctgggcagg
tggagatggc cagtgatgtc caataaagga cactggaggg agcagtgtga gtaaaggccc
                                                                                  160260
                                                                                  160320
tgagggcatt catgttcagg gagggttgct gcccactggc ttgcttggca cacaggagag
tgggtattcc tgccttagta actttatgta aacaagtatt tcctcagtct gttcctctca
                                                                                  160380
                                                                                  160440
                                                                                  160500
aactgootgo totggéacat toagaatgto acagaactea cotggatgea tteagecoot
160560
gtcagactgg cccagtctgt gggcaaggag cctagagagg gcttagtttc agcttgaaag
                                                                                   160620
gagctgggat ttaccaagaa gcaaatgaga gacgaggatt gcaacaactg tgccatttcc ccagcttcag ctgactcctg tatattgact gtgccttcag actcatccgt aagtgacccc
                                                                                   160680
                                                                                  160740
aggetgget eteceacate acagtaagaa ttecacacae catacaactt ggaaagagge tecagetgaa ggaageeca caettette aagttttet tagtettete ttettggeaa agagtaeett ttgttette taattatgta actattggtt tagtaaatat teaeceatte
                                                                                  160800
                                                                                   160860
                                                                                  160920
agtcaccetg taagtggcag gcactgttta cagggacaca ggaaggaata aaaacttgca
ggcaccttgg agcttgcatt ctattgaaga ggtaatggaa gttgggatag cagctaaact
                                                                                   160980
                                                                                  161040
atgetggtat tggccaggeg cagtggetca cacetgtaat eccageaett tggaggecaa
                                                                                   161100
ggtgggcaga tcatgaagtc aggagatcga gaccatcctg gctaacatgg tgaaaccccg tctctactaa aagtaaaaaa aaaaattagc caggtgtggt ggcgggcgcc tgtagtccca
                                                                                   161160
                                                                                   161220
                                                                                   161280
gctacttggg aggctgaggc aggagaatgg tgtgaaccca ggaggcgaag attgcagtga
                                                                                   161340
gccgagatgg caccactgca ctccagcctg ggtgacagag cgagactctg tctcagaaaa
aaaaaatatg ctggtagttt tgattcaaga tggcctttgg agcccatgat ttaggtctcg
                                                                                   161400
tacccaccaa ggtctactgg aaaacatcag gctctcctgc tatagaccca tagggagagc
                                                                                   161460
tgcagccgag agggggagct gaagagaagt gccccttctg tgtcctgtca gcctcatcct
                                                                                   161520
                                                                                   161580
teegeaagga ceagttgetg tgecacteea tteacttget geaagactgg aggtttttee
traggtgttg agracetggt tracaagatg tragratett gatgertgag accatraagg
                                                                                   161640
caagictotg aacagggott accttagagt aaggottaga agaggoogta aagtcagtot
                                                                                   161700
cageteegtg getetgeaga getttgggae atgtgaatte ttaaaaacaa gaetattgta cagttactat atgeatgeag tataaaatta taacettgga aaateetage tagetgttga
                                                                                   161760
                                                                                   161820
                                                                                   161880
gctaattcca taaagtaatc agctcctgag ttctgcagtg gtaataataa tcagcataat gagtaaacac tgtgtgtgcc aggcagcgtc tcatttgatc cttgtgataa tcttgtaagt
                                                                                   161940
actgattttc toccttottt aaacaaagtt ttttttttt ttttagagag ggtotcacta
                                                                                   162000
                                                                                   162025
tgttgcccag gctagtcttg aattc
<210> 37
<211> 1350
<212> DNA
<213> Homo Sapien
<220>
<221> CDS
<222> (213)...(920)
<308> GenBank AJ242973
<309> 1999-10-26
<400> 37
                                                                                       60
geggeeget egaegtgaea geeggtaege eegggtttgg geaacetega ttaegggegg
celecaggee egecageage geocegegee geoegeeege geceetgeeg ecceeeggtt
                                                                                     120
coggoogogg accocactot obgoogttoo ggotgoggot cogotgoogg tagogoogto
                                                                                      180
ccccgggacc accettcgge tggcgccete ce atg ete teg gee ace egg agg
Met Leu Ser Ala Thr Arg Arg
                                                                                      233
get tgc cag etc etc etc etc cac age etc ttt ecc gte ecg agg atg
                                                                                      281
Ala Cys Gln Leu Leu Leu His Ser Leu Phe Pro Val Pro Arg Met
```

		10					15		٠			20					
ggc Gly	aac Asn 25	tcg Ser	gcc Ala	tcg Ser	aac Asn	atc Ile 30	gtc Val	agc Ser	ccc Pro	cag Gln	gag Glu 35	gcc Ala	ttg Leu	ccg Pro	ggc Gly		329
cgg Arg 40	aag Lys	gaa Glu	cag Gln	acc Thr	cct Pro 45	gta Val	gcg Ala	gcc Ala	aaa Lys	cat His 50	cat His	gtc Val	aat Asn	ggc Gly	aac Asn - 55		377
aga Arg	aca Thr	gtc Val	gaa Glu	cct Pro 60	ttc Phe	cca Pro	gag Glu	gga Gly	aca Thr 65	cag Gln	atg Met	gct Ala	gta Val	ttt Phe 70	gga Gly		425
atg Met	gga Gly	tgt Cys	ttc Phe 75	tgg Trp	gga Gly	gct Ala	gaa Glu	agg Arg 80	aaa Lys	ttc Phe	tgg Trp	gtc Val	ttg Leu 85	aaa Lys	gga Gly		473
gtg Val	tat Tyr	tca Ser 90	act Thr	caa Gln	gtt Val	ggt Gly	ttt Phe 95	gca Ala	gga Gly	ggc Gly	tat Tyr	act Thr 100	tca Ser	aat Asn	cct Pro	•	521
act Thr	tat Tyr 105	aaa Lys	gaa Glu	gtc Val	tgc Cys	tca Ser 110	gaa Glu	aaa Lys	act Thr	ggc Gly	cat His 115	gca Ala	gaa Glu	gtc Val	gtc Val		569
cga Arg 120	gtg Val	gtg Val	tac Tyr	cag Gln	cca Pro 125	gaa Glu	cac His	atg Met	agt Ser	ttt Phe 130	gag Glu	gaa Glu	ctg Leu	ctc Leu	aag Lys 135		617
gtc Val	ttc Phe	tgg Trp	gag Glu	aat Asn 140	cac His	gac Asp	ccg Pro	acc Thr	caa Gln 145	ggt Gly	atg Met	cgc Arg	cag Gln	999 Gly 150	aac Asn		665
gac Asp	cat His	ggc Gly	act Thr 155	cag Gln	tac Tyr	cgc Arg	tcg Ser	gcc Ala 160	atc Ile	tac Tyr	ccg Pro	acc Thr	tct Ser 165	gcc Ala	aag Lys		713
caa Gln	atg Met	gag Glu 170	gca Ala	gcc Ala	ctg Leu	agc Ser	tcc Ser 175	aaa Lys	gag Glu	aac Asn	Tyr	caa Gln 180	aag Lys	gtt Val	ctt Leu		761
			ggc Gly														809
			tat Tyr														857
ccc Pro	aat Asn	ggc Gly	tac Tyr	tgc Cys 220	ggc Gly	ctt Leu	Gly 999	ggc Gly	acc Thr 225	ggc Gly	gtg Val	tcc Ser	tgc Cys	cca Pro 230	gtg Val		905
			aaa Lys 235		ttg	ctcc	cca (catg	gtgg	gc c	tttg	aggt	t cc	agta	aaaa		960
															tgcac cgcga		1020 108 <u>,</u> 0

```
1140
ggcaagttga taaaatgtga cttatctcct aataagttat ggtgggagtg gagctgtgcg
gtttcctgtg tettetgggg tetgagtgaa gatageaggg atgetgtgtt caccettett
                                                                          1200
ggtagaaget aaggtgtgag etgggaggtt getggacagg atgggggace ceagaagtee tttatetgtg etetetgee gecagtgeet tacaatttge aaacgtgtat ageetcagtg
                                                                          1260
                                                                          1320
                                                                          1350
acteattege tgaaateett egetttacea
<210> 38
<211> 235
<212> PRT
<213> Homo Sapien
<400> 38
Met Leu Ser Ala Thr Arg Arg Ala Cys Gln Leu Leu Leu His Ser
Leu Phe Pro Val Pro Arg Met Gly Asn Ser Ala Ser Asn Ile Val Ser
             20
Pro Gln Glu Ala Leu Pro Gly Arg Lys Glu Gln Thr Pro Val Ala Ala
                                                     45
                              40
Lys His His Val Asn Gly Asn Arg Thr Val Glu Pro Phe Pro Glu Gly
                                                60
Thr Gln Met Ala Val Phe Gly Met Gly Cys Phe Trp Gly Ala Glu Arg
                                            75
                      70
Lys Phe Trp Val Leu Lys Gly Val Tyr Ser Thr Gln Val Gly Phe Ala
                  85
Gly Gly Tyr Thr Ser Asn Pro Thr Tyr Lys Glu Val Cys Ser Glu Lys
                                   105
             100
Thr Gly His Ala Glu Val Val Arg Val Val Tyr Gln Pro Glu His Met
         115
                               120
                                                    125
Ser Phe Glu Glu Leu Leu Lys Val Phe Trp Glu Asn His Asp Pro Thr
                          135
    130
Gln Gly Met Arg Gln Gly Asn Asp His Gly Thr Gln Tyr Arg Ser Ala
                                                                  160
                      150
                                           155
Ile Tyr Pro Thr Ser Ala Lys Gln Met Glu Ala Ala Leu Ser Ser Lys
                                       170
                  165
Glu Asn Tyr Gln Lys Val Leu Ser Glu His Gly Phe Gly Pro Ile Thr
                                                         190
             180
                                   185
Thr Asp Ile Arg Glu Gly Gln Thr Phe Tyr Tyr Ala Glu Asp Tyr His
                               200
                                                     205
         195 -
Gln Gln Tyr Leu Ser Lys Asn Pro Asn Gly Tyr Cys Gly Leu Gly Gly
                                                220
                          215
Thr Gly Val Ser Cys Pro Val Gly Ile Lys Lys
                                            235
<210> 39
<211> 481
<212> DNA
<213> Homo Sapien
<300>
<308> GenBank AW195104
<309> 1999-11-29
<400> 39
ggcattattg gactgtaggt ttttattaaa acaaacattt ctcatagctc taagcaaagc
                                                                              60
attagaatto atcaagogga ctcacatott ttototgcac agagaggggc tgaaaaggga
                                                                             120
gagaaagtcc cttatgtatg tctagatttg gtaaagcgaa ggatttcagc gaatgagtca
                                                                             180
ctgaggctat acacgtttgc aaattgtaag gcactggcgg gcagagagca cagataaagg acttctgggg tcccccatcc tgtccagcaa cctcccagct cacaccttag cttctaccaa
                                                                             240
                                                                             300
gaagggtgaa cacagcatee etgetatett cacteagace ceagaaaace cagggaaace
                                                                             360
 cgacagetee acteccacca taaettatta ggagataagt cacattttat caaettgeca
                                                                             <u>4</u>20
```

tegegee g	ctcc	tatagattat	acttcggtaa	acccaatctg	tataaattcc	tttgtacttt	480 481
<210> 4 <211> 3 <212> D <213> H	390 ONA	Sapien					
<300> <308> 0 <309> 2		nk AW874187 05-22	7			· .	
gcattag agagaaa actgagg gactttt caagaag	ttat gaat agcc gcta tggg.	tcatcaagcg ccttatgtat tacacgtttg ggtccccat	gactcacatc gtctagattt caaattgtaa tcctgtccag atccctgcta	ttttctctgc ggtaaagcga ggcactggcg caacctccca	tteteatage acagagaggg aggattteag ggcagagage geteacacet gacceccaga	ctgaaaaggg cgaatgagtc acagataaag tagcttctac	60 120 180 240 300 360 390
<210> 4 <211> 4 <212> I <213> F	43 ONA	ficial Seque	ence			·	
<220> <223> 0	Oligo	onucleotide	Primer				
<400> 4 agcggat		aatttcacac	agggagctag	cttggaagat	tgc	•	43
<210> 4 <211> 2 <212> I <213> 3	22 DNA	ficial Seque	ence				
<220> <223> 0	Oligo	onucleotide	Primer				
<400> 4 gtccaat		tgcaaacagt	tg				22
<210> 4 <211> 2 <212> I <213> J	23 DNA	ficial Sequ	ence				
<220> <223> (Oligo	onucleotide	Primer				
<400> 4 agcggat		aatttcacac	agg				23
<210 > 4 <211 > 3 <212 > 1 <213 > 3	18 DNA	ficial Sequ	ence				
<220>				•		٠.	^

<223>	Oligonucleotide Primer	
<400> actgag	44 geetg etgeataa	18
<210><211><211><212><213>	21	
<220> <223>	Oligonucleotide Primer	
<400> tctcaa	45 atcat gtgcattgag g	21
<210><211><211><212><213>	43	
<220> <223>	Oligonucleotide Primer	
<400> agcgg	46 ataac aatttcacac agggatcaca cagccatcag cag	43
<210><211><211><212><213>	23	
<400> agcgg	47 ataac aatttcacac agg	23
<210><211><211><212><213>	18	
<400> ctggc	48 gccac gtggtcaa	18
<210><211><211><212><213>	20	
<220> <223>	Oligonucleotide Primer	
<400>	49 ctgca cagagagggc	20
<210><211><212><213>	44	
<220>		_

<223>	Oligonucleotide	Primer					
<400> agcgga	50 Itaac aatttcacac	agggctgaaa	teettegett	tacc			44
<210><211><211><212><212>	23	ence		•			
<220> <223>	Oligonucleotide	Primer	,				
<400> agcgga	51 staac aatttcacac	agg					.23
<210><211><211><212><213>	18	ence					
	Oligonucleotide	Primer			·		
<400> ctgaaa	52 aggg agagaaag						18
<210><211><211><212><213>	20	ence					
<220> <223>	Oligonucleotide	Primer					•
<400> tcccaa	53 aagtg ctggaattac					٠.	20
<210><211><211><212><213>	22	ence					
<220> <223>	Oligonucleotide	Primer					
<400> gtccaa	54 atata tgcaaacagt	tg					22
<210><211><211><212><213>	20	ence					
<220> <223>	Oligonucleotide	Primer					
<400> cccaca	55 agcag ttaatccttc						20

<210><211><212><213>	18		
<220> <223>	Oligonucleotide primer		
<400> gcgcto	56 cetgt eggtgeea		18
<210><211><211><212><213>	18		
<220> <223>	Oligonucleotide primer	•	·
<400> gcctga	57 actgg tggggccc		18
<210><211><212><213>	15		
<220> <223>	Oligonucleotide primer		
<400> catgca	58 atgca eggte		15
<210> <211> <212> <213>	30		
<220> <223>	Oligonucleotide primer	·	
<400> cagaga	59 agtac ccctcgaccg tgcatgcatg		30
<210><211><211><212><213>	15		
<220> <223>	Oligonucleotide primer		
<400> catgc	60 atgca cggtt		1!
<210><211><211><212><213>	30		

<220> <223>	Oligonucleotide primer				
<400> gtacgt	61 caegt gecaacteee catgagagae				30
<210><211><212><212><213>	14				
<220> <223>	Oligonucleotide primer	•			
<400> catgc	62 atgca cggt				14
<210><211><212><213>	18			• .	
<220> <223>	Oligonucleotide primer				
<400> gcctg	63 actgg tggggccc				18
<210><211><211><212><213>	26				
<220> <223>	Oligonucleotide primer				
<400> gtgct	64 gcagg tgtaaacttg taccag				26
<210><211><211><212><213>	28				
<220> <223>	Oligonucleotide primer				
<400> cacgg	65 atccg gtagcagcgg tagagttg				28
<210><211><212><212><213>	19	·			
<220> <223>	Oligonucleotide primer				
<400>	66 gcatg tggagacag	•			ر1:

	·	
<210>	67	
<211>	18	
<212>	DNA	
	Artificial Sequence	
<220>		
<223>	Oligonucleotide primer	
<400>	67	
	tctt gccatgag	18
	3	
<210>	68	
<211>		
<212>		
<213>	Artificial Sequence	
224	•	
<220>	Oligonucleotide primer	
<223>	Oligonacieotide primer	
<400>	68	
	acga cgtt	14
-		
<210>		
<211>		
<212>		
<213>	Artificial Sequence	
<220>	· ·	
	Oligonucleotide primer	
~	Oligonadicolia primor	
<400>	69	
cggata	acaa tttc	1
<210>		
<211>		
<212>	Artificial Sequence	
(213)	Artificial bequence	
<220>		
	Oligonucleotide primer	
<400>		
caattt	catc gctggatgca atctgggcta tgagatc	3
<210>	71 -	
<211>		
<212>		
	Artificial Sequence	
<220>		
<223>	Oligonucleotide primer	
-400:		
<400>		3
caatit	caca cageggatge ttettttgge tetgaet	,
<210>	72	
<211>		
<212>		
	Artificial Sequence	

<220> <223>	Oligonucleotide primer			
<400> tcagto	72 cacga cgttggatgc caataaaagt	gactctcagc		40
<210><211><211>	37	``````````````````````````````````````		
	Artificial Sequence			
<220> <223>	Oligonucleotide primer			
<400>				
cggata	acaa tttcggatgc actgggagca	ttgaggc		37
<210>			•	
<211>				
<212> <213>	Artificial Sequence			. •
<220>				
<223>	Oligonucleotide primer			
<400>	74			
tcagt	cacga cgttggatga gcagatccct	ggacaggc		38
<210>		•		
<211>				
<212><213>	Artificial Sequence			
<220>	•			
	Oligonucleotide primer			
<400>		. *		
cggata	aacaa tttcggatgg acaaaatacc	tgtattcc		38
<210>				
<211>				
<212>	Artificial Sequence			
	•			
<220> <223>	Oligonucleotide primer			
<400>	76			
tċagto	cacga cgttggatgc agagcagctc	cgagtc		36
<210>				•
<211>				
<212><213>	DNA Artificial Sequence		_	
<220> <223>	Oligonucleotide primer			
<400>	77			
	gtgat cattggatgc aggaagctct	gg	,	, 32

	·		
<210><211><211>	38 DNA		
<213>	Artificial Sequence		
<220> <223>	Oligonucleotide primer		
<400> tcagt	78 cacga cgttggatgc ccacatgcca cccactac	•	. 36
<210> <211> <212> <213>	35		
<220> <223>	Oligonucleotide primer		
<400> cggata	79 aacaa ttteggatge eegteaggta eeaeg		35
<210><211><212><213>	37		
<220> <223>	Oligonucleotide primer		
<400> tcagto	80 cacga cgttggatgc ccacagtgga gcttcag		37
<210>	81		
<211>	22		
<212>	DNA Artificial Sequence		
			, ,
<220> <223>	Oligonucleotide primer	٠.	
<400>	81		
gctcat	tacct tgcaggatga cg	,	23
<210>			
<211>			
<212>	Artificial Sequence		
(2137	Attitional bequence		
<220> <223>	Oligonucleotide primer		
<400>			_
tcagto	cacga cgttggatga ccagctgttc gtgttc	٠	3
<210>			
<211>	· ·		
<212>	Artificial Sequence		
	ten en e		

<220> <223>	Oligonucleotide primer	
<400> tacat	83 ggagt teggggatge acaeggegae tete	34
<210> <211> <212> <213>	40	
<220> <223>	Oligonucleotide primer	
<400>	84	
tcagt	cacga cgttggatgg ggaagagcag agatatacgt	40
<210>		
<211><212>		•
	Artificial Sequence	
<220>		
	Oligonucleotide primer	
<400>	85	
gaggg	gctga tccaggatgg gtgctccac	29
<210>	86	
<211>		
<212> <213>	Artificial Sequence	
<220>		
	Oligonucleotide primer	
<400>		20
tgaag	cactt gaaggatgag ggtgtctgcg	30
<210>		
<211><212>		•
	Artificial Sequence	
<220>		
	Oligonucleotide primer	
<400>	87	
cggat	aacaa tttcggatgc tgcgtgatga tgaaatcg	38
<210>		
<211><212>		
	Artificial Sequence	•
<220>		
	Oligonucleotide primer	
<400>	88	
	agete deaggatgee agagge	26

<210><211><212><213>	27	
<220> <223>	Oligonucleotide primer	
<400> gccgcc	89 cggtg taggatgctg ctggtgc	27
<210><211><211><212><213>	31	
<220> <223>	Oligonucleotide Template	-
<400> cgcagg	90 ggttt cctcgtcgca ctgggcatgt g	31
<210><211><212><212><213>	43	
<220> <223>	Biotinylatd primer	
<400> tgctta	91 atoco tgtagotaco otgtottggo ottgoagato caa	43
<210> <211> <212> <213>	42	
<220> <223>	Oligonucleotide primer	
<400> agcgga	92 ataac aatttcacac aggecateac acegeggtac tg	42
<210><211><211><212><213>	44	
<220> <223>	Oligonucleotide primer	
<400> cccagt	93 tcacg acgttgtaaa acgtcttggc cttgcagatc caag	44
<210><211><212>	42	

<220> <223>	Oligo	nucleotide	primer					
<400> agcgga		aatttcacac	aggccatcac	accgcggtac	tg			42
<210><211><211><212><213>	20 DNA	icial Seque	ence					
<220> <223>	Oligo	onucleotide	primer	÷ .				
<400> ctccag		gcaggagtgc						20
<210><211><211><212><213>	17 DNA	icial Seque	ence				•	
<220> <223>	Oligo	onucleotide	primer					
<400>		gctccct						17
<210> <211> <212> <213>	23 DNA	icial Seque	ence					
<220>. <223>		nylated pri	mer		·	•		
<400> cccagt		acgttgtaaa	acg	·		*		23
<210><211><211><212><213>	100 DNA	sapien						•
	gagaa	agggctctgc agtcagagcc			ctgcaacaat	ctgggctatg		60 100
<210><211><211><212><213>	100 DNA	sapien					,	
	gagaa	agggctctgc agtcagagcc			ctgcaacaat	ctgggctatg	·	60 100
<210> <211> <212>	100	-						

			,			
<213> Homo	sapien					
<400> 100 gaattatttt aatgctccca	tgtgtttcta gtgctattca	aaactatggt tgggcagctc	tcccaataaa tctgggctca	agtgactctc	agcgagcctc	60 100
<210> 101 <211> 100 <212> DNA <213> Homo	sapien		·			
<400> 101 gaattatttt aatgctccca	tgtgtttcta gtgctattca	aaactatggt tgggcagctc	tcccaataaa tctgggctca	agtgactctc	agcaagcctc	60 100
<210> 102 <211> 84 <212> DNA <213> Homo	sapien	•	٠.,			
<400> 102 taataggact tttgtccttg	acttctaatc aagtaacctt	tgtaagagca tcag	gatecetgga	caggcgagga	atacaggtat	60 . 84
<210> 103 <211> 84 <212> DNA <213> Homo	sapien				* .	
	acttctaatc aagtaacctt		gatecetgga	caggcaagga	atacaggtat	60 84
<210>-104 <211> 100 <212> DNA <213> Homo	sapien			·		
	gcatttgatt ctgtgggcat		teegagteeg atgtetegta	tccagagctt	cctgcagtca	60 100
 <210> 105 <211> 100 <212> DNA <213> Homo	sapien		·			•
	gcatttgatt ctgtgggcat		teegagteea atgtetegta	tccagagctt	cctgcagtca	60 100
<210> 106 <211> 100 <212> DNA <213> Homo	sapien					٠.
	cctgcaaggg cgggcatcgt		ggcccacatg cagggctgcg	ccacccacta	ccagggcacg	60 100

			,			
<210> 107 <211> 100 <212> DNA <213> Homo sa	apien		, ,	. •		•
<400> 107 agcaaggact co tggtacctga co	tgcaaggg gggcatcgt	ggacagtgga cagctggggc	ggcccacatg cagggctgcg	ccacccacta	ccggggcacg	60 100
<210> 108 <211> 100 <212> DNA <213> Hom sap	oien					
<400> 108 caataactct aa ccggggcgtc aa	atgcagcgg acctgcaag	aagatgacct gtatgagcat	gcccacagtg accccccttc	gagcttcagg	gcgtggtgcc	60 100
<210> 109 <211> 100 <212> DNA <213> Homo sa	npien					
<400> 109 caataactct aa ccggggcgtc aa				gagcttcagg	gcttggtgcc	60 100
<210> 110 <211> 100 <212> DNA <213> Homo sa	apien	•				1.
<400> 110 ttgaagettt gg gtgtggagee ee	ggctacgtg cgaactcca	gatgaccagc tgggtttcca	tgttcgtgtt gtagaatttc	ctatgatcat	gagagtegee	60 100
<210> 111 <211> 100 <212> DNA <213> Homo sa	apien					
<400> 111 ttgaagcttt gg gtgtggagcc co				ctatgatgat	gagagtcgcc	60 100
<210> 112 <211> 100 <212> DNA <213> Homo sa	apien				•	
<400> 112 ggataacctt gg caggcctgga to				tacgtgccag	gtggagcacc	60 100
<210> 113 <211> 100 <212> DNA <213> Homo sa	apien	• .				
<400> 113						. •

ggataacctt caggcctgga	ggctgtaccc tcagcccctc	cctggggaag attgtgatct	agcagagata gggagccctc	tacgtaccag	gtggagcacc	60 100
<210> 114 <211> 80 <212> DNA <213> Homo	sapien	÷				
<400> 114 tgaagcactt ttgaggctga	gaaggagaag cacattcttc	gtgtctgcgg	gagccgattt	catcatcacg	cagcttttct	60 80
<210> 115 <211> 80 <212> DNA <213> Homo	sapien					
	gaaggagaag cacattcttc	gtgtctgcgg	gagtcgattt	catcatcacg	cagcttttct	60 80
<210> 116 <211> 80 <212> DNA <213> Homo	sapien			. •		
	gctcccagaa gcggcccctg	tgccagaggc	tgctccccgc	gtggcccctg	caccagcagc	60 80
<210> 117 <211> 80 <212> DNA <213> Homo	sapien		·			
	gctcccagaa gcggcccctg	tgccagaggc	tgctccccc	gtggcccctg	caccagcagc	60 80
<210> 118 <211> 48 <212> DNA	ficial Seque	ence			·	
<220> <223> Hair	pin structu	ıre				
<400> 118 cagagagtac	ccctcaaccg	tgcatgcatg	aaacatgcat	gcacggtt		4.8

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 19 April 2001 (19.04.2001)

(26) Publication Language:

(30) Priority Data:

60/159,176

60/217,658

60/217,251

09/663,968

PCT

English

US

US

US

(10) International Publication Number WO 01/027857 A3

(51) International Patent Classification7:	G06F 19/00	US Filed on
(21) International Application Number:	PCT/US00/28413	US Filed on
(22) International Filing Date: 13 October	2000 (13.10.2000)	(71) Applicant (for all
(25) Filing Language:	English	QUENOM, INC. San Diego, CA 921

13 October 1999 (13.10.1999)

19 September 2000 (19.09.2000)

10 July 2000 (10.07.2000)

10 July 2000 (10.07.2000)

(72) Inventors; and

ll designated States except US): SE-[US/US]; 3595 John Hopkins Court, 121 (US).

(75) Inventors/Applicants (for US only): BRAUN, Andreas [DE/US]; 11237-6 Carmel Creek Road, San Diego, CA 92130 (US). KÖSTER, Hubert [DE/CH]; Via Delle Scuole 1, CH-6900 Lugano-Cassarate (CH). VAN DEN BOOM, Dirk [DE/DE]; Eppendorfer Weg 205 D, D-20253 Hamburg (DE). PING, Yip [US/US]; 3641 Copley Avenue, San Diego, CA 92116 (US). RODI, Charlie [US/US]; 13823 Recuerdo Drive, Del Mar, CA 92014 (US). HE, Liyan [CN/US]; 10948 Creek Bridge Place, San Diego, CA 92128 (US). CHIU, Norman [CA/US]; 1128 Caminito Alvarez, San Diego, CA 92126 (US). JURINKE, Christian [DE/DE]; Rombergstrasse 22, 20255 Hamburg (DE).

[Continued on next page]

09/663,968 (CIP)

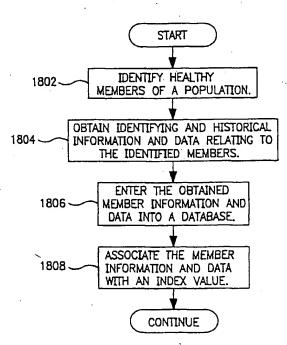
60/217,251 (CIP) 10 July 2000 (10.07.2000)

19 September 2000 (19.09.2000)

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 60/159,176 (CIP) Filed on 13 October 1999 (13.10.1999) US 60/217,658 (CIP) 10 July 2000 (10.07.2000) Filed on

(54) Title: METHODS FOR GENERATING DATABASES AND DATABASES FOR IDENTIFYING POLYMORPHIC GENETIC MARKERS



(57) Abstract: Process and methods for creating a database of genomic samples from healthy human donors, methods that use the database to identify and correlate polymorphic genetic markers and other markers with diseases and conditions are provided.

WO 01/027857 A3

- (74) Agents: SEIDMAN, Stephanie, L. et al.; Heller Ehrman White & McAuliffe, Suite 700, 4250 Executive Square, La Jolla, CA 92037 (US).
- (81) Designated States (national): AE, AG, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- (88) Date of publication of the international search report: 3 October 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

Intern: 1 Application No PCT/US 00/28413

	·	PC"	T/US 00/	28413			
A. CLASSIF	FICATION OF SUBJECT MATTER G06F19/00						
1PC /	900L TA\ 00	•					
According to	International Patent Classification (IPC) or to both national classification	lion and IPC					
B. FIELDS		n symbols)					
IPC 7	currentation searched (classification system followed by classification $G06F$	ii Syntholaj					
Documentati	ion searched other than minimum documentation to the extent that su	ch documents are included i	in the fields sea	ırched			
Electronic de	ala base consulted during the international search (name of data base	e and, where practical, searc	ch terms used)				
WPI Dat	ta, EPO-Internal						
	•	-					
-	•						
C. DOCUME	ENTS CONSIDERED TO BE RELEVANT						
Category •	Citation of document, with indication, where appropriate, of the rele	vant passages		Relevant to clai	im No.		
x	WO 99 05323 A (AFFYMETRIX INC)	`		1-100			
r l	4 February 1999 (1999-02-04)	-	. 1	•			
	abstract; claims 1,4	•					
X	WO 97 40462 A (SPECTRA BIOMEDICAL	INC)		1-100			
ļ	30 October 1997 (1997-10-30)		1				
·	page 4, line 2 - line 22 page 7, line 3 -page 8, line 4						
ļ			}				
Х	WO 98 24935 A (AN GANG ; HARA MARK	0 (US);	ļ	1-100			
	RALPH DAVID (US); VELTRI ROBERT (11 June 1998 (1998-06-11)	us); U)					
	page 4, line 27 -page 5, line 6	-		•			
٠	page 6, line 14 - line 18				÷		
		/					
,	,						
				•			
		[C] n	nom 255 71 11				
	her documents are listed in the continuation of box C.	χ Patent family memb	uers are listed it	n annex.			
	tegories of cited documents :	"T" later document published or priority date and not i	in conflict with t	he application but			
consid	ent defining the general state of the art which is not dered to be of particular relevance	cited to understand the invention	principle or the	ory underlying the			
	"E" earlier document but published on or after the international filling date "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to						
"L" docume	"L' document which may throw doubts on priority claim(s) or hydrow an inventive step when the document is taken alone which is cited to establish the publication date of another "Y" document of particular relevance; the claimed invention						
"O" docume	n or other special reason (as specified) ent referring to an oral disclosure, use, exhibition or	cannot be considered to document is combined	o involve an inve with one or mor	entive step when the re other such docu-			
other r	means	ments, such combination in the art.	-				
later th	nan the priority date claimed	"&" document member of the in					
Date of the	actual completion of the international search	Date of malling of the in	~	. J., . open			
ì	0 September 2001	28/09/2001	<u> </u>				
Name and n	mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2	Authorized officer		^			
	European Patent Office, P.B. 5818 Patentiaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo ni,	E211 ^					
	Fax: (+31-70) 340-3016	Filloy Gar	Cia, E				

INTERNATIONAL SEARCH REPORT

Intern: Il Application No
PCT/US 00/28413

		PCT/US 00/28413
C.(Continu	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category *	Citation of document, with Indication, where appropriate, of the relevant passages	Relevant to claim No.
X	COLLINS F S ET AL: "A DNA Polymorphism Discovery Resource for Research on Human Genetic Variation" GENOME RESEARCH, vol. 8, 1998, pages 1229-1231, XP002177106 the whole document	1-100
P,X	WO 00 51053 A (BRYANT STEPHEN PAUL ;GEMINI RESEARCH LTD (GB); KELLY PAUL JAMES (G) 31 August 2000 (2000-08-31) abstract; claims 1-24	1-100
(WO 98 35609 A (HELMS RONALD W ;TOMASKO LISA (US); BIOMAR INTERNATIONAL INC (US);) 20 August 1998 (1998-08-20) abstract; claims 1-25	1-100
,	SARKAR C ET AL: "Human Genetic Bi-allelic Sequences (HGBASE), a Database of Intra-genic Polymorphisms" MEM INST OSWALDO CRUZ, 'Online! vol. 93, no. 5, September 1998 (1998-09) - October 1998 (1998-10), pages 693-694, XP002177107 Rio de Janeiro Retrieved from the Internet: <url:http: brook="" cgb="" es="" groups="" publications.htm="" www.cgb.ki.se=""> 'retrieved on 2001-09-05! the whole document</url:http:>	1-100
. 1	FOSTER M W AND FREEMAN W L: "Naming Names in Human Genetic Variation Research" GENOME RESEARCH, vol. 8, 1998, pages 755-757, XP002177108 the whole document	2,9,27, 55,72
		·
	-	
, ,		

INTERNATIONAL SEARCH REPORT

Inten al Application No PCT/US 00/28413

Patent document cited in search report			Publication date			Publication date	
WO 990	5323	A	04-02-1999	EP EP EP WO WO US	1002264 A 1009861 A 0998697 A 1007737 A 9905574 A 9905324 A 9905591 A 6229911 B 6188783 B	24-05-2000 21-06-2000 10-05-2000 14-06-2000 04-02-1999 04-02-1999 04-02-1999 08-05-2001 13-02-2001	
WO 974	0462	A	30-10-1997	AU EP JP	2734197 A 0897567 A 2000508912 T	12-11-1997 24-02-1999 18-07-2000	
WO 982	4935	Α	11-06-1998	AU AU EP US	722819 B 5515198 A 0960214 A 6190857 B	10-08-2000 29-06-1998 01-12-1999 20-02-2001	
WO 005	1053	Α	31-08-2000	AU	2815900 A	14-09-2000	
WO 983	5609	Α	20-08-1998	US AU BR CN EP	6059724 A 6151498 A 9807366 A 1268033 T 0973435 A	09-05-2000 08-09-1998 18-04-2000 27-09-2000 26-01-2000	